Symptom2Vec 을 활용한 병력 청취 시스템

김민지 [1], 조인휘 [1]
[1] 한양대학교 컴퓨터소프트웨어학과
a2927036@hanyang.ac.kr, iwjoe@ hanyang.ac.kr

A medical history taking system using Symptom2Vec

Min-Ji Kim[1], In-whee Jee[1]
[1]Dept. of Computer Software, Hanyang University

요 약

임상 환경에서 진료시간의 대부분은 환자의 증상을 듣고, 추가 증상을 이끌어내는데 사용된다. 이를 병력 청취라고 하며, 진료에 있어서 가장 기본적이고 필수적인 활동이다. 하지만 병력 청취에 대한 연구가 1940 년대부터 계속되고 있음에도 아직까지 표준이 정립되지 않았으며, 다양한 분야에 접목되는 딥러닝 기술 또한 병력 청취와 관련해서는 연구가 부족한 현실이다. 본 논문에서는 Symptom2Vec 을 새롭게 제안하였으며, 이를 활용하여 질병에 따른 증상의 평균 cosine 유사도 점수(0.962)로 병력 청취의 기준을 확립하였다. 또한 most similar word Top5 를 확인하는 것을 통해 환자의 증상에 따른 유사 증상을 묻는 병력 청취가 가능한 것을 확인하였다. 이를 통해 실제 임상 환경에서의 자동화된 병력 청취 시스템을 제안한다.

1. 서론

진료에 있어서 병력 청취는 환자의 경험을 토대로 의사의 진단 정확도를 크게 높일 수 있는 활동으로, 반드시 기초적으로 수행해야한다. 병력 청취를 이용해 수집한 환자의 증상은 질병 진단에 있어서 결정적인 단서가 되기 때문이다. 예를 들면 어지럼은 흔한 증상 중 하나지만, 의식소실이 함께 있다면 중추성 또는 심장질환의 가능성을 시사하기 때문에 어지럼 증상이 있는 환자에게 체계적이고 자세한 병력 청취는 필수적이다. [1]

따라서 본 연구는 병력 청취를 자동화할 수 있는 시스템을 연구하였다. 현재 미리 수집된 데이터를 활용하여 머신러닝을 학습하고, 이를 통해 질병을 진단하고 예측하는 연구는 많이 진행되고 있다. [2] 하지만 인공지능 분야에서 병력 청취와 관련된 연구는 부족한 상황이다. 본 연구에서는 병력 청취의 자동화를 위하여 Word2Vec 을 기반으로 한 Symptom2Vec 을 새롭게 제안한다.

2. 데이터셋

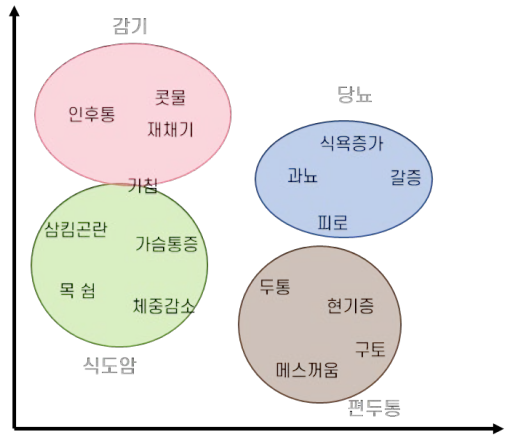
데이터는 미국과 영국의 공인 데이터인 WebMD Dictionary, NHS inform, snomed ct, clevelandclinic 등을 이용하여 수집했으며, 각 사이트에서 수집한 데이터의 중복을 제거하는 등의 전처리 과정을 진행했다. 이를 통해 총 526 개의 질병과 2615 개의 증상이 포함된 데이터셋을 완성하였다. 해당 데이터셋의 구성은 <표 1>과 같다.

No.	Attribute	Description
1	Name	질병명
2	Symptom	질병에 해당하는 증상
3	Accompany	해당 질병의 동반질환

<표 1> 데이터셋 구성

3. 제안 방법

본 논문에서는 효율적인 병력 청취를 위하여 자연어처리 분야에서 주로 사용되는 Word2Vec 에 기반한 Symptom2Vec 을 새롭게 제안하였다. 이는 기존 Word2Vec 과 다르게 문장에 해당하는 단어가 아닌, 질병에 해당하는 증상을 vector 상에 embedding 한다. 이를 통해 Symptom2Vec 은 그림 1 과 같이 특정 질병에 해당하는 증상끼리 군집이 형성될 수 있다.



(그림 1) Symptom2Vec 예시

이는 환자의 증상과 유사도가 높은 증상을 묻는 것을 통해 추가 문진을 가능하게 하며, 일정 군집이 형성되는지 여부에 따라 추가문진의 진행 여부도 판단할 수 있다.

질병에 따른 증상의 연관성과 추가문진의 진행여부를 파악하기 위해서는 군집 형성이 잘 되는지 계산이 필요하다. 이는 각 질병의 증상 리스트로 정의한다. 해당 기준은 질병에 따른 증상들의 유사도 점수(cosine similarity score)를 각각 계산하여 평균을 구해 수식 1 과 같이 계산한다.

기준 유사도 점수 = (질병 1 의 증상 유사도 평균 + 질병 2 의 증상 유사도 평균 + … + 질병 526 의 증상 유사도 평균) / 526 (1)

질병의 증상 유사도 평균은 증상 간의 cosine similarity score 를 구해 평균을 낸 것이며, 이렇게 생성한 기준 유사도 점수를 활용하여 추가 문진을 진행할 것인지에 대한 유무를 판별할 수 있다.

4. 실험 및 결과

본 논문에서는 데이터셋의 symptom 과 accompany 를 합쳐 질병에 해당하는 증상들의 집합을 Skip-gram 모델을 활용하여 학습하였다. 학습된 Symptom2Vec 모델에 증상을 넣었을 때, 유사 증상은 <표 2>와 같다.

Cough	유사도	Headache	유사도
Sore throat	0.981	chills	0.989
Muscle aches	0.978	Muscle aches	0.985
Headache	0.975	fever	0.979
Rash	0.974	Couch	0.975
Chest pain	0.970	Rash	0.972

<표 2> Symptom2Vec 에 Cough 와 Headache 를 입력했을 때, 유사 증상 Top 5 와 유사도

증상 유사도 평균	유사도
최소값	0.855
최대값	0.995
평균	0.962

<표 3> 수식 1 로 계산한 Sympom2Vec 의 증상 유사도 평균

<표 2>를 통해 Symptom2Vec 이 환자의 증상이 입력될 경우, 유사 증상을 묻는 과정에서 병력 청취가 가능한 것을 확인할 수 있다. <표 3>은 수식 1 을 활용한 계산 결과를 정리한 것으로, 질병에 따른 증상들의 유사도 평균 최소값이 0.855 인 것을 확인했을 때, 군집이 잘 형성된 것을 알 수 있다. 또한 모든 질병의 증상 유사도 평균을 기준으로 추가 문진에 대한 기준을 정의할 수 있다.

5. 결론

본 논문에서는 질병에 따른 증상을 vector 상에 표현하여 Symptom2Vec 을 통한 병력 청취 시스템 연구를 진행하였다. 그 결과, 질병에 따른 증상의 평균 cosine 유사도 점수가 0.962 로 확인되었고 최소 유사도 점수는 0.855 로 확인되었다. 이를 통해 증상의 군집도를 파악할 수 있었으며, 병력 청취 수집 지속성에 대한 기준을 확립할 수 있었다. 그리고 Symptom2Vec 에 임의의 증상을 넣어 most similar word Top5 를 확인하는 과정을 통해 환자의 증상에 따른 유사 증상을 확인할 수 있었다.

이를 활용할 경우 임상의가 환자의 증상을 이끌어내는 시간을 줄이고, 환자가 공개하기 꺼려하는 정보에 있어서 환자 순응도를 높여 정보 수집이 원활하게 이루어질 수 있을 것이라 예상한다. 실제로 [3]의 연구에서 PDA 방식을 사용할 경우, 93%의 환자 순응도가 보고되었다.

또한 증상을 토대로 질병을 유추하고, 해당 질병에 대한 증상을 추가 문진을 진행하는 연역법적 문진 시스템은 특정 질병에 대한 추론 편향 문제를 야기할 수 있지만, Symptom2Vec 은 환자의 증상을 토대로 유사 증상을 수집하는 과정을 통해 귀납법적 문진을 진행하기 때문에 추론 편향을 예방할 수 있어 실제 임상 환경에서 활용도가 클 것이라 예상한다.

향후 Symptom2Vec 과 담화분석 모델을 결합하여 실시간으로 병력 청취를 진행할 수 있는 의료용 챗봇 연구를 진행할 예정이다. 이를 통해 환자에게 질병 및 진료과를 추천할 경우, 기존의 방법보다 편의성과 정확도 측면에서 높은 성능의 추천이 가능할 것이라고 기대하는 바이다.

참고문헌

1. Ban, J. and M. Kim, *Art of history taking in dizzy patients.* Research in Vestibular Science2012. 11: p. 3-5.
2. Ahsan, M.M., S.A. Luna, and Z. Siddique. *Machine-Learning-Based Disease Diagnosis: A Comprehensive Review*. in *Healthcare*. 2022. MDPI.
3. Dale, O. and K.B. Hagen, *Despite technical problems personal digital assistants outperform pen and paper when collecting patient diary data.* Journal of clinical epidemiology, 2007. 60(1): p. 8-17.