

경량 딥러닝과 지식베이스를 활용한 모바일 질환별 식품 추천 시스템

현범수*, 김도현**, 이상근***

*고려대학교 컴퓨터학과

**고려대학교 컴퓨터·전파통신공학과

***고려대학교 인공지능학과

bshyeon@korea.ac.kr, dhkim1028@korea.ac.kr, yalphy@korea.ac.kr

Mobile Food Recommendation System for Patients Using Light-weight Deep Learning and Knowledge Bases

Bumsu Hyeon*, Dohyun Kim**, SangKeun Lee***

*Dept. of Computer Science and Engineering, Korea University

**Dept. of Computer and Radio Communications Engineering, Korea University

***Dept. of Artificial Intelligence, Korea University

요 약

본 논문에서는 딥러닝과 지식베이스를 융합하여 활용한 질환 인식 및 식품 추천 시스템을 제안한다. 제안하는 시스템은 온전히 모바일 디바이스 내에서 작동하는 시스템이다. 본 시스템은 압축된 딥러닝 모델을 이용해 사용자 대화 텍스트를 분석하여 사용자의 질환을 예측한다. 그 후, 지식베이스를 기반으로 해당 질환 관리에 도움이 되는 식품을 매칭하고 사용자에게 추천한다. 이는 사용자 친화적 헬스케어 애플리케이션으로써 체크리스트 작성 등 번거로운 작업 없이도 사용자에게 유용한 건강 정보를 제공할 수 있다.

1. 서론

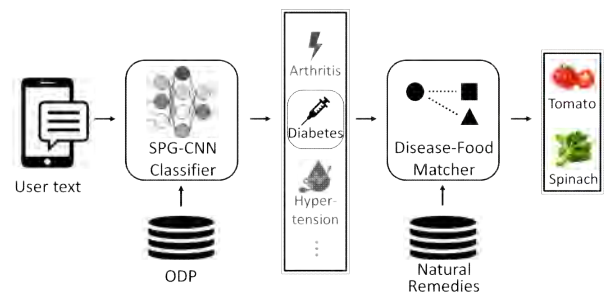
고령화와 만성 질환의 만연으로 스마트 헬스케어의 수요가 증가하고 있다. 이에 따라 딥러닝을 활용한 다양한 헬스케어 응용 시스템이 제안되었고, 괄목할 만한 성능 및 효용성을 보이고 있다[1]. 또한 관련 분야에서 딥러닝과 지식베이스를 융합한 시스템에 관한 연구 역시 활발하게 이루어지고 있다.

하지만 헬스케어 애플리케이션을 구축하기 위해 서버-클라이언트 구조를 사용하면 민감정보로 분류되는 건강정보를 취급하는 데 있어 네트워크 전송, 저장 등에 엄격한 규제가 적용된다. 반면, 시스템 전체를 사용자의 디바이스에서 구동할 경우, 이러한 문제를 피할 수 있으나 컴퓨팅 자원이 현저히 제한된다.

본 논문에서는 위와 같은 문제를 해결하기 위해 사용자 디바이스 내에서 구동되는 경량 헬스케어 시스템을 제안한다. 본 시스템은 모바일 환경에서 구동되며, 사용자에게 개인화된 질환별 식품 추천 서비스를 제공한다. 본 시스템에서는 경량 딥러닝 모델인 SPG-CNN[2]을 채택하고, 딥러닝 모델 압축

기법을 적용하여 매우 적은 리소스를 사용하는 텍스트 분류 모델을 구축한다. 텍스트 분류 모델을 이용해 사용자 텍스트로부터 질환을 예측한 후 지식베이스 Natural Remedies를 활용해 예측된 질환과 추천 후보 식품을 매칭하여 사용자에게 제공한다.

제안하는 시스템은 질환 예측 정확도 66.7%로 상당한 예측률을 나타내어 만성 질환자 및 위험군 사용자들에게 건강 관리를 위한 유용한 정보를 제공할 수 있을 것으로 기대된다.



(그림 1) 시스템 구조.

2. 제안하는 시스템

그림 1은 전체 시스템 구조를 보여준다. 본 시스템은 모바일 디바이스의 메신저 대화 텍스트를 추출

한 후 텍스트 분류기를 통해 대화 정보로부터 사용자가 가질 수 있는 질환을 예측한다. 예측된 질환은 질환-식품 매칭을 통해 식품과 매칭된다.

2.1 질환 예측

질환 예측에는 딥러닝 모델 SPG-CNN[2] 및 지식베이스 ODP¹⁾를 사용한다. SPG-CNN은 경량 합성곱 신경망(convolutional neural network)으로, 멀티태스크 학습(multi-task learning) 기법을 활용한 텍스트 분류 모델이다. 본 시스템에서는 SPG-CNN을 8비트 양자화(8-bit quantization)를 통해 압축하여 사용한다. 또한 SPG-CNN의 파라미터 중 상당수를 차지하는 워드 임베딩(word embedding)에는 압축률을 높이기 위해 별도로 Compositional Code Learning 압축 기법[3]을 사용한다. 표 1에서 위 압축 기법 사용 시 모델 정확도에 거의 손실을 주지 않고 9배 이상의 압축률을 달성함을 확인할 수 있다. SPG-CNN의 학습 데이터에는 ODP를 사용한다. ODP는 온라인 참여자들이 400만여 개의 웹페이지를 주제에 따라 80만여 개의 카테고리로 분류한 지식베이스이다. 본 시스템에서는 [4]에서 제안한 방법으로 ODP를 정제한 후 질환명에 해당하는 카테고리 그 카테고리에 속하는 웹 페이지를 학습 데이터로 사용한다.

	크기	정확도
압축 전	233 MB	67.2%
압축 후	25 MB	66.7%

(표 1) 압축 전·후 모델 크기 및 정확도

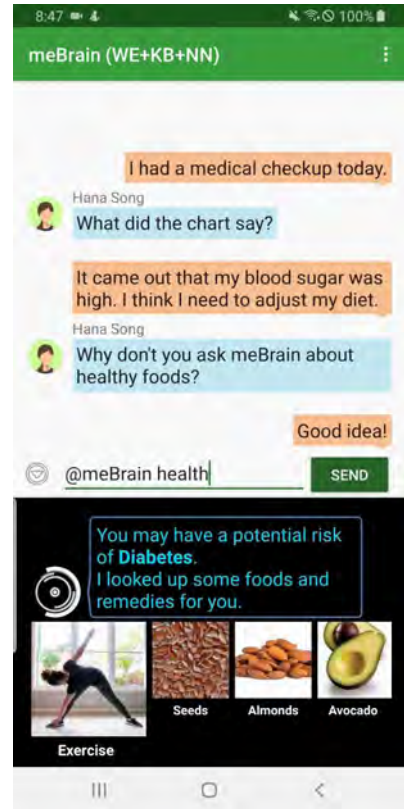
2.2 질환-식품 매칭

질환과 식품을 매칭하는 과정에는 지식베이스 Natural Remedies²⁾를 룩업하는 방식을 사용한다. Natural Remedies는 질환별 추천 식품을 추천도 순으로 정리한 지식베이스로, 121개의 질환에 대응되는 485개의 식품이 존재한다. 여기에는 식품 외에도 운동 등 생활습관 관련 내용도 일부 포함되어 있다. 본 시스템에서는 추천도가 높은 4개의 항목을 사용자에게 제시한다. 그림 2는 안드로이드 스마트폰에서 실행 중인 본 시스템을 보여준다.

3. 결론

본 논문에서는 압축된 딥러닝 모델과 지식베이스

를 활용하여 모바일 환경에서 사용자의 대화 텍스트로부터 질환을 예측하고, 해당 질환의 관리에 도움이 되는 식품을 추천하는 시스템을 제안하였다. 향후 본 시스템은 예측된 질환 정보를 활용하여 의료 정보 등 더 다양한 서비스를 제공하는 플랫폼으로 확장될 수 있을 것으로 기대된다.



(그림 2) 질환 예측 및 식품 추천 화면.

참고문헌

- [1] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep Learning for Healthcare: Review, Opportunities and Challenges," Briefings in Bioinformatics, 2017, pp. 1 - 11.
- [2] K. Kim, Y. Kim, J. Lee, J. Lee, and S. Lee, "From Small-scale to Large-scale Text Classification," The World Wide Web Conference (WWW), San Francisco, 2019, pp. 853-86.
- [3] R. Shu and H. Nakayama, "Compressing Word Embeddings via Deep Compositional Code Learning," International Conference on Learning Representations (ICLR), Vancouver, 2019.
- [4] J. Lee, J. Ha, J. Jung, and S. Lee, "Semantic Contextual Advertising Based on the Open Directory Project," ACM Trans. Web (TWB), vol. 7, no. 4, 2013, article 24.

1) <http://odp.org/>

2) <http://classic.personalremedies.com/Remedy.asp>