

챗봇을 이용한 강아지 건강관리 전문가 시스템

임지수*, 김다영*, 조수민*, 유건아*

*덕성여자대학교 컴퓨터학과

e-mail:da1599@naver.com

Dog's Health Care Management Expert System Using Chatbot

Ji-Su Im*, Da-Yeong Kim*, Su-Min Cho* ,
Kyeon-Ah Yu*

*Dept of Computer Science, Duk-Sung Women's University

요 약

반려동물 인구 천만 시대에 맞는 맞춤형 의료 서비스의 필요성 증대 등 국내 반려동물 시장의 성장세에 맞추어 다양한 반려동물 관련 앱 서비스가 출시되고 있다. 그중 의료 관련 분야의 서비스의 수요가 고령 반려동물의 증가에 따라 특히 급속도로 증가하는 분야임에도 불구하고 관련 사례나 연구가 미약하다. 따라서 본 논문에서는 기존에 제공되는 서비스 중 추론의 비효율성 문제를 극복하고 추론 과정에서의 사용자 피드백을 활용하여 궁극적으로 진단의 효율성을 극대화할 수 있는 전문가 시스템을 제시하였다.

1. 서론

최근 반려동물을 키우는 가정이 많아지면서 반려동물과 관련된 많은 서비스가 제공되고 있다. 특히 반려동물을 가족처럼 생각하는 인식이 커지면서 반려동물 의료 서비스의 수요가 늘어나게 되고 그에 따른 비용 부담도 크다. 반려동물과 함께하는 많은 사람들이 이런 비용 부담을 줄이기 위해 반려견의 증상을 인터넷으로 검색한 후 직접 약국에서 약을 구입해 먹이는 경우도 적지 않다. 또 반려동물 특성상 병원에 직접 데려가야 하기 때문에 질병에 걸린 반려견의 응급처치에 대한 대처가 늦어질 수 있다. 마지막으로 인터넷의 정확하지 않은 정보 중 사용자가 필요한 정보를 찾아내는 중간 과정에서 소요되는 시간과 정보의 신뢰성에 대한 문제가 있다.

이를 해결하기 위한 기존의 시스템으로는 반려동물을 키우는 사람들을 위한 국내 온라인 커뮤니티 서비스인 ‘아지냥이’의 챗봇 기반의 진단 서비스[1]와 웹으로 강아지 증상에 대한 질병 정보를 제공하는 ‘PET MD’의 Dog Symptom Checker 서비스[2] 등이 있다. 하지만 이러한 시스템 중 아직까지 자연어로 증상을 입력받아 처리되는 한국어 전용 시스템은 보고된 바 없으며 설명 시스템 또한 갖춰지지 않아 추론의 근거를 설명할 수 없다. 이러한 문제를 해결하고자 본 논문에서는 챗봇을 기반으로 반려견의 건강과 관련된 지식을 제공하는 건강관리 전문가 시스템을 제안하였다. 제안하는 시스템에서는 자연어로 증상을 입력받아 질병을 추론하는 추론엔진이 존재하며 추론 과정에서 사용자의 입력에 따라 지식베이스가 업데이트되는데 이는 다음 [3]사용자가 세부 증상을 선택하는데 도움을

주며 이전 사용자의 피드백을 토대로 진단 결과의 신뢰도 정보를 제공하여 진단에 참고할 수 있다. 질병 추론이 완료되었을 때 해당 진단의 근거를 제공하는 설명 시스템 또한 갖추고 있으며 설명 시스템을 위해서 챗봇이 직접 수행하는 동사 부분의 제한적인 자연어 처리 프로세스가 존재한다. 또한 해당 시스템은 반려견의 질병뿐 만 아니라 건강관리의 전반적인 측면에서 정보를 제공한다. 먼저 동물 전용 약품명으로 약품에 대한 정보를 검색할 수 있으며 견종에 따라 자주 걸리는 질병 정보 및 이에 대한 상세정보도 조회할 수 있다.

2. 챗봇 기반의 전문가 시스템

2.1 전문가 시스템(Expert System)

전문가 시스템이란 컴퓨터가 어떤 특정 분야의 전문가 처럼 판단을 하여 전문적인 정보를 제공함으로써 비전문가도 전문적인 지식에 접근할 수 있도록 하는 시스템이다. 시스템 구축을 위해 사용된 챗봇은 사용자의 입력에 따라 채팅 형태로 대화를 할 수 있는 사용자 인터페이스(User Interface)로 응답의 주체는 챗봇이며 사용자가 선택한 기능에 따라 설정된 다이얼로그로 들어가며 대화가 진행된다. 챗봇의 응답의 바탕이 되는 정보는 전문가로부터 습득된 지식이며 이는 적절한 형태로 구조화되어 사용자에게 제공되는 정보로 활용되는데 이것을 지식베이스(Knowledge Base)라고 한다. 이 지식베이스에서 특정 메커니즘을 거쳐 사용자가 원하는 정보를 찾아내도록 하는 것이 추론엔진(Inference Engine)이다. 마지막으로 추론한

지식에 대한 근거를 제공하는 것이 설명 시스템(Explanation System)이며 이 4가지가 전문가 시스템을 구성하는 기본 요소라고 할 수 있다.

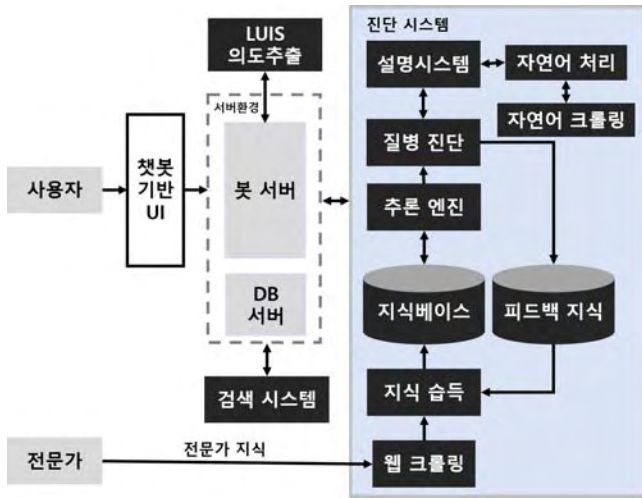


그림 1. 시스템 구성도

2.2 사용자 인터페이스(User Interface)

사용자는 웹·앱 기반의 챗봇 UI를 통해 전문가 시스템에 접근할 수 있다. 추론을 하기 위한 전처리 단계에서는 사용자에게 추론에 사용되는 정보를 입력받으며 LUIS가 의도를 분류하여 추론엔진으로 넘긴다.

2.3 추론 엔진(Inference Engine)

챗봇은 사용자로부터 받은 자연어로 된 문장에서 의도(Intent)를 분류하는데 의도가 증상(Symptom)이라면 해당 문장에서 개체(Entity)를 추출하면서 추론이 시작된다. 예를 들어 사용자가 입력한 문장이 '설사를 해요'라면 의도는 '간단한 질병'이며 개체는 '설사'이다. 추론을 위한 첫 번째 세부 증상(Specific Symptom)이 설사로 결정되었으므로 지식베이스에서 설사와 관련된 두 번째 세부 증상의 리스트를 불러온다. 여기서 사용자는 해당하는 세부 증상을 선택하게 되고 이에 대한 세 번째 세부 증상이 존재한다면 해당 세부 증상 리스트를 불러온 후 세 번째 세부 증상을 입력받을 준비를 한다. 추론엔진은 이 과정을 반복하며 증상에 따른 질병을 진단한다.

2.4 지식베이스(Knowledge Base)



그림 2. 강아지 증상·질병 지식베이스 구조

지식베이스는 추론에 이용되는 전문가로부터 습득한

지식으로 웹에서 읽어와 CSV 파일로 저장된 내용을 이용하였다. 질병 추론을 위한 지식베이스는 세부 증상들(symptom, specific_symptom)과 해당 세부 증상에 대한 질병명(disease), 질병에 대한 상세정보를 볼 수 있는 URL과 간단한 대처법(handle), 증상 출현도(ss_num)와 질병 신뢰도(select_num)로 구성하였다. 증상 출현도는 이전 사용자가 가장 많이 선택한 순서대로 각 증상의 항목에 우선순위를 부여하여 보여 준다. 따라서 증상 출현도는 사용자에게 세부 증상 선택의 편의성을 제공하며 선택지가 과부하되는 것을 방지할 수 있다. 질병 신뢰도란 추론한 각 증상들이 이전 사용자들에게 얼마나 신뢰를 얻었는가 하는 정도이다. 추론엔진을 거쳐 진단된 질병은 하나일 수도 있고 그 이상일 수도 있다. 하나라면 해당 증상에 대한 신뢰도가 무조건 100%가 되는 것이 아니라 사용자의 피드백에 따라 신뢰도를 전혀 얻지 못할 수도 있다. 진단한 질병이 여러 개일 경우 신뢰도가 더 높은 질병이 우선순위를 가지게 되고 사용자는 진단에 이를 참고할 수 있다. 따라서 추론엔진이 한 번의 추론 과정을 거칠 때마다 지식베이스가 업데이트되고 다음 사용자에게 즉시 반영되며 이를 피드백 지식이라고 한다.

2.5 설명 시스템(Explanation System)

설명 시스템은 추론엔진이 어떻게 추론했는가를 나타내는 단계별 사고과정을 설명하는 부분이다. 즉 진단 시스템에서 설명 시스템은 사용자에게 진단한 질병에 대한 추론의 근거를 설명한다. 진단 결과를 받은 사용자들이 어떻게 이런 결과가 도출되었는지에 대한 근거를 요청하면 사용자가 입력한 각 단계의 세부 증상들이 모두 설명 시스템이 제공하는 설명에 사용된다. 이때 동사로 끝나는 여러 개의 세부 증상 문장들을 하나의 자연스러운 문장으로 만들기 위해 자연어 처리가 이루어지고 이 하나의 문장이 추론의 근거가 된다.

3. 구현 및 테스트

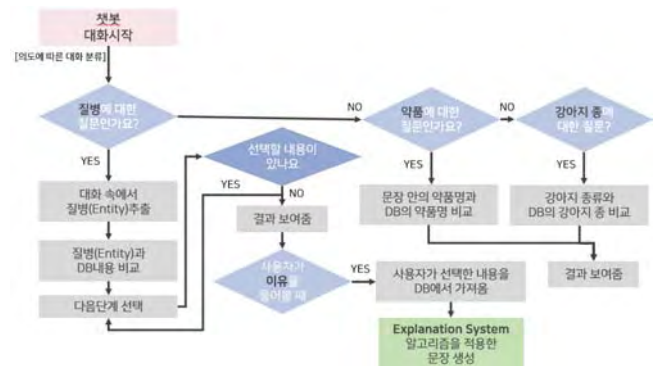


그림 3. 챗봇 알고리즘

3.1 증상 관련 서비스

사용자가 서비스 메뉴에서 증상 서비스를 선택하면 챗봇은 반려견의 증상에 대해 묻는다. 사용자가 증상에 대해

말하면 챗봇은 사용자가 말한 문장의 의도(Intent)를 파악하여 각 상황에 따른 대답을 해야 하고 필요할 경우 사용자의 문장 내에서 증상 정보에 해당하는 개체(Entity)를 추출해야 한다. 이 부분은 Microsoft에서 제공하는 LUIS 자연어 처리 기능을 사용했다. LUIS를 통해 추출된 증상 정보는 지식베이스의 증상 데이터와 비교된다. 추출된 증상 정보가 증상 데이터에 존재하면 사용자가 선택한 증상보다 더 좁은 범위의 세부 증상들을 리스트로 보여주고 사용자의 선택을 기다린다. 사용자가 세부 증상을 선택할 때 선택한 세부 증상을 포함한 지식베이스의 증상 데이터의 ss1_num 항목은 1씩 증가한다. 이는 출현도를 계산하는데 사용된다. 사용자가 세부 증상을 선택한 뒤에 챗봇은 사용자가 선택한 세부 증상보다 더 좁은 범위의 세부 증상 데이터를 리스트로 사용자에게 보여주고 사용자의 선택을 기다리는 작업을 반복한다.



그림 4. 세부 증상 입력 화면

이러한 과정은 다음 세부 증상이 존재할 때까지 진행된다. 사용자가 가장 좁은 단위의 세부 증상을 선택하면 챗봇은 증상에 해당하는 질병과 대처법들을 리스트로 안내하며 최종 진단을 내린다. 진단 리스트 중 사용자가 특정 질병명을 선택했을 때 해당 진단에 대한 평가를 내릴 수 있다. 도움이 되었다고 응답했을 경우에는 해당 진단을 포함하는 지식베이스의 질병 데이터의 select_num 항목이 1씩 증가하는데 이는 신뢰도를 계산하는데 사용된다.



그림 5. 질병 추론 결과 화면

설명 시스템은 지식베이스의 여러 문장을 한 문장으로 자연스럽게 연결하기 위한 자연어 처리 알고리즘을 사용해 구현한다. 먼저, 지식베이스의 문장에서 동사나 형용사를 추출하고 C# 크롤러를 이용해 해당 단어를 네이버 국어 사전에 검색해 원형을 크롤링해 가져온다. 가져온 단어들

의 어미를 원하는 형태로 변형하여 여러 문장들을 하나의 완벽한 문장으로 사용자에게 진단의 이유를 설명한다.

3.2 검색 관련 서비스

약품 검색과 견종에 따른 질병 검색 서비스는 비슷한 형식으로 진행된다. 먼저 사용자가 서비스 메뉴에서 약품 혹은 견종에 관한 질병 검색 서비스를 선택하면 챗봇은 약품명 또는 견종을 입력하도록 유도한다. 사용자가 검색 내용을 입력하면 챗봇은 입력 결과와 약품 지식베이스[4]의 내용을 비교한다. 지식베이스에 검색 내용이 포함되어 있을 경우 포함된 모든 내용이 카드리스트로 표현된다. 약품 검색인 경우, 약품의 종류와 약품명이 표시되고 해당 약품을 클릭할 때 더 자세한 정보가 있는 링크로 연결된다. 견종 검색인 경우, 견종과 해당 견종이 잘 걸리는 질병이 나타나고 해당 견종을 클릭하면 해당 견종이 잘 걸리는 질병에 대한 정보링크로 연결된다.

4. 결론

챗봇을 이용한 강아지 건강관리 전문가시스템은 사용자가 직접 증상을 검색해 원하는 정보를 찾는 시간을 줄여 주므로 시간적 측면에서 효율적이다. 또 비전문가도 전문성을 요구하는 분야에 손쉽게 접근할 수 있다는 측면에서 경제적으로도 효율적인 서비스라고 할 수 있다.

기존의 동물 진단 서비스 중 자연어로 받은 증상을 처리하는 기능이 없어서 해당하는 질병을 사용자가 일일이 찾아야 한다. 현재 해당 서비스의 첫 번째 세부 증상은 24+a개(기타 증상 포함)이기 때문에 진단에 더 많은 시간이 걸린다. 첫 번째 세부 증상만 걸려내도 다음 증상이 10개 미만으로 줄어들기 때문에 데이터가 같다고 하면 제안하는 시스템이 더 빠르게 진단이 가능하다. 또 챗봇을 이용한 강아지 건강관리 전문가 시스템에서는 사용자의 피드백을 기반으로 더 자주 발생한 증상의 우선순위를 높이기 때문에 사용자 피드백 데이터가 쌓일수록 사용자가 리스트의 원하는 세부 증상을 빠른 시간 안에 찾을 수 있도록 한다. 진단한 질병 리스트의 순서도 사용자 피드백을 기반으로 보여주기 때문에 질병 예측이 더 수월해질 수 있다.

참고문헌

[1] 아지냥이, <https://www.samsungcard.com/personal/mobile/pet/UHPPMO0114M0.jsp>, 검색일 2018.9
 [2] PET MD, https://www.petmd.com/symptom-checker?icn=Tools_SuBNav&icl=1_symptom_checker, 검색일 2018.9
 [3] 코구레 노리오, <https://terms.naver.com/list.nhn?cid=58401&categoryId=58401>, 검색일 2017.10
 [4] 대한동물약국협회, <http://anipharm.net/anidict>, 검색일 2018.3