

음성인식 텍스트 분석을 통한 자동 일정 관리 시스템 개발에 관한 연구

이해미¹, 조위덕²

¹아주대학교 전자공학과

²아주대학교 전자공학과

dlgoal1327@ajou.ac.kr, chowd@ajou.ac.kr

A Study on the Development of Automatic Schedule Management System through Speech Recognition Text Analysis

Hae-Mi Lee¹, We-Duke Cho²

¹Dept. of Electronic Engineering, Ajou University

²Dept. of Electronic Engineering, Ajou University

요 약

컴퓨터가 마이크 등의 소리 센서를 통해 얻은 음향학적 신호를 단어나 문장으로 변환시키는 기술인 음성 인식 기술과 인공지능 기술을 결합한 음성 대화 시스템에 대한 연구 진행 및 제품 출시가 활발하게 이루어지고 있다. 기존의 시스템을 사용하면서 날짜와 시간 외의 정보 추출 정도가 빈약하거나 자동 등록이 되지 않는 문제점을 확인하였다. 음성 인식 기술을 통해 얻은 텍스트에서 보다 많은 정보를 추출하고, 자동 등록 및 알림과 맛집 등 추가 정보 제공 시스템을 구축하는 것을 목표로 하였다.

1. 서론

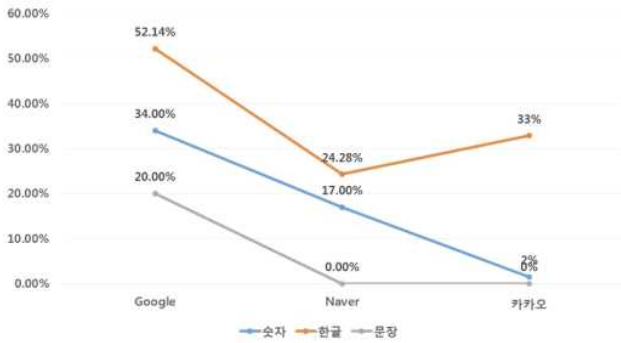
음성 인식 기술은 컴퓨터가 마이크 등의 소리 센서를 통해 얻은 음향학적 신호를 단어나 문장으로 변환시키는 기술이다[1]. 이와 인공지능 기술을 결합한 음성 대화 시스템에 대한 연구 진행 및 제품 출시가 활발하게 이루어지고 있다. 스마트폰, AI 스피커 등에 요구사항을 전달하였을 때, 이를 처리해주는 기능은 계속해서 발전되어 왔다. 그러나 해당 기능을 사용하여 일정을 등록하고자 할 때, 날짜와 시간 외의 정보 추출 정도가 빈약하거나 자동 등록이 되지 않는 문제점을 발견할 수 있었다. 이에 따라 음성 인식 기술을 통해 얻은 텍스트에서 보다 많은 정보를 추출하고, 자동 등록 및 알림 시스템을 구축하는 것을 목표로 하였다.

본 연구는 룰베이스로 음성인식을 통해 얻은 텍스트에서 추출한 정보를 기반으로 일정을 등록해주는 시스템 구축에 대한 것이다. 음성인식 API로는 구글과 카카오를 사용하였고, 이를 통해 얻은 텍스트 데이터에서 룰베이스로 정보를 추출하였다. 이후엔 해당 정보를 바탕으로 구글 캘린더에 일정을 등록하였으며, 등록 직후에 카카오톡으로 일정에 대한

정보를 정리하여 전송하도록 하였다. 추가적으로 매일 오전 8시마다 당일의 일정에 대한 정보를 전송하도록 하였으며, 장소에 대한 정보가 있는 경우에는 크롤링을 통해 얻은 맛집 정보를 정리하여 제공하도록 하였다.

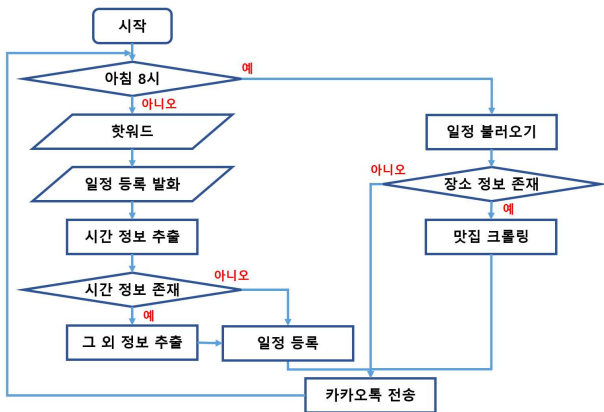
2. 관련연구

음성인식을 통한 일정 관리와 관련된 연구는 다음과 같은 연구들이 있었다. 먼저 남지순(2009)은 숫자를 포함한 시간 표현을 추출하는 기술에 대한 연구를 진행하였다[2]. 해당 연구에서는 텍스트에서 관찰 가능한 시간 표현의 다양한 언어적 형태와, 이를 기반으로 한 하위 유형화에 대해 다루었다. 다음으로 최승주 외 1명(2017)은 음성 API에 관련하여 정확도를 분석하였다[3]. 해당 논문에서는 Google, Naver, 카카오 세 종류의 음성인식 API에 대해 숫자, 한글, 문장의 인식률에 대한 결과를 확인할 수 있었다. 중점으로 확인한 숫자와 문장의 정확도가 가장 높은 API는 카카오였고, 한글에 대해서는 Naver의 정확도가 가장 높은 것을 확인할 수 있었다.



(그림 1) 참고문헌 [3]의 음성인식 API 실험 결과

3. 시스템 구조

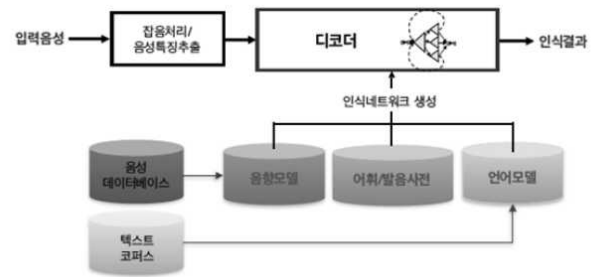


(그림 2) 시스템 Flow Chart

시스템 Flow Chart를 통해 시스템 동작 방식에 대해 알 수 있다. 먼저 프로그램이 실행되면 현재 시각이 아침 8시인지 확인한다. 하루 일정을 시작하기 전, 당일의 모든 일정 정보를 제공하기 위함으로 시간은 사용자가 설정 가능하다. 아침 8시로 판단하게 되면 구글 캘린더 API를 사용하여 당일의 일정을 모두 불러오고, 각 일정별로 장소 정보를 확인하여 존재하면 맛집을 크롤링한다. 그 이후엔 일정에 대한 정보를 카카오톡으로 전송하게 된다.

일정 등록을 위한 발화를 하기 전, 핫워드를 인식하는 과정을 추가하여 일상에서 오인식되지 않도록 하였다. 만약 핫워드가 인식이 되면 STT를 통해 일정 등록에 대한 발화를 문장으로 변환시킨다. 이렇게 얻은 문장에서 시간 정보 추출을 시도하고, 시간이 추출되면 사람, 장소, 일정명 등의 정보를 추출한다. 그 이후엔 추출된 정보를 기반으로 일정을 등록하고, 이에 대한 정보를 카카오톡으로 전송하게 된다.

4. Speech-To-Text (STT)



(그림 3) ASR (Automatic Speech Recognition) System 구성 [4]

음성인식은 STT(Speech-To-Text)라고도 불리는데, 키보드나 기타 입력 장치에 의한 입력이 아니라 사람의 음성을 통한 입력이 되므로 인공지능 비서, 챗봇 등 다양한 기술의 기반이 된다. STT는 발화자의 음성을 기계적인 알고리즘을 통해 텍스트로 변환을 수행한다[5]. STT의 성능은 DB 크기와 품질에 비례하여 향상될 수 있다. 상용 서비스에 적용되는 음향 모델은 대부분 확률 통계 방식인 HMM(Hidden Markov Model) 기반으로 이루어졌다. 최근에는 시퀀스-투-시퀀스(Sequence-to-Sequence) 방식의 RNN(Recurrent Neural Network) 기반으로 속도와 성능 면에서 좋은 결과를 가져오기도 했다[6].

참고문헌을 바탕으로 종합적인 정확도가 높았던 카카오 STT를 사용하기로 하였다. 그러나 음성인식의 시작을 파악하기 위해 설정한 핫워드를 인식하는데에 다소 딜레이가 있는 것으로 보였다. 구글 STT의 경우 파이썬에서 API 사용을 위한 인증 토큰을 발행하지 않고 사용 가능했기 때문에 두 가지의 경우에 대한 인식 시간을 비교하였다. 동일한 음성 과일을 준비하고 실행시킨 결과 구글 STT의 경우엔 약 0.36초, 카카오 STT의 경우엔 약 0.71초로 거의 2배 정도 차이가 나는 것을 확인할 수 있었다. 이러한 이유로 핫워드 인식에는 구글 STT를 사용하였으며, 정확한 정보 추출을 위해 비교적 정확도가 높아야 하는 일정 등록 문장 인식에는 참고 문헌 [3]에서 정확도가 가장 높았던 카카오 STT를 사용하게 되었다.

5. 텍스트 정보 추출

처음 텍스트에서 정보를 추출하고자 했을 때, 룰베이스가 아닌 개체명 인식을 사용하고자 하였다.

그러나 생각만큼 정확도가 높지 않아, 정확한 정보 추출을 목표로한 연구에 맞지 않는 방법이라 판단하였다. 이에 따라 해당 시스템에서 사용되는 문장의 특징에 대해 생각해보게 되었다. 일정 등록을 위해 전달하는 말에 담긴 정보의 양은 한정적이며, 비교적 짧은 문장으로 구상된다. 이러한 문장의 특징을 이용하여 덤러닝이 아닌 룰베이스로 텍스트에서 정보를 추출하게 되었다. 본 연구는 다양한 기능을 제공하는 것이 아닌 일정 등록 기능 특화 시스템을 구축하는 것이 목표였으므로, 모든 발화는 일정 등록에 관련되었다고 가정한 후 진행하였다.

추출하고자 했던 정보는 날짜, 시간, 사람, 장소, 일정명으로 총 5가지이다.

<표 1> 사람, 장소, 일정명 추출 Rule

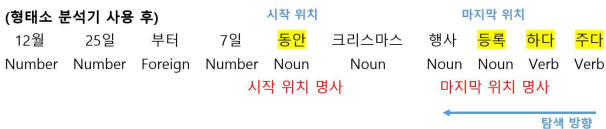
사람	~랑, ~이랑' 등
장소	~에서' 등
일정명	문장 끝부분의 조사, 특정 명사와 동사, 형용사, 특정 명사 사이의 명사

먼저 사람, 장소, 일정명에 대한 추출 방법이다. 사람 정보 추출 시에는 '랑'이라는 조사가 붙는 어절을 선택한 후 조사를 제거하고, 끝에 '이'가 붙는 경우 받침 여부를 확인하여 제거하는 방식을 사용하였다. 이 외에도 '~와', '~과' 등 구어체, 문어체에 따른 다른 표현 방식에 대해서도 조사를 제거하는 방식을 사용하였다. 장소 정보 추출 시에는 '에서'라는 조사가 붙는 어절을 선택하여 조사를 제거하는 방식을 사용하였다.

(원문장)

12월 25일부터 7일 동안 크리스마스 행사 등록해 줘

(형태소 분석기 사용 후)



(그림 4) 일정명 추출 방식

마지막으로 일정명 정보 추출 시에는 문장의 특성을 고려하였다. 일정을 정의할 시에 어떤 것에 대한 일정인지에 대한 정보는 보통 끝에 위치한다. 이에 따라 형태소 분석기를 사용하여 문장을 형태소로 분리하고, 문장 끝부분부터 탐색하여 마지막으로 사용된 조사, 동사, 형용사의 위치를 확인한다. 그 후

엔 명사 중 '동안' 등 다른 정보의 나열을 끝마치는 명사는 시작 위치 명사로, '등록' 등 문장을 끝마치는 명사는 마지막 위치 명사로 정의한다. 이를 바탕으로 조사, 시작 위치 명사와 동사, 형용사, 마지막 위치 명사 사이의 명사를 찾아 일정명으로 등록하게 되는 것이다. 예를 들어 설명하자면 '12월 25일부터 7일 동안 크리스마스 행사 등록해 줘'라는 문장이 있다면, '동안'부터 '등록', '하다' 사이의 명사인 '크리스마스 행사'를 일정명으로 추출하게 되는 것이다.

<표 2> 날짜, 시간 추출 Rule

	절대적	상대적
날짜	~월', '~일', '~째주 ~요일'	~주 뒤', '~일 후', '~요일' 등
시간	~시', ~분'	~시', '~시간', ~분'

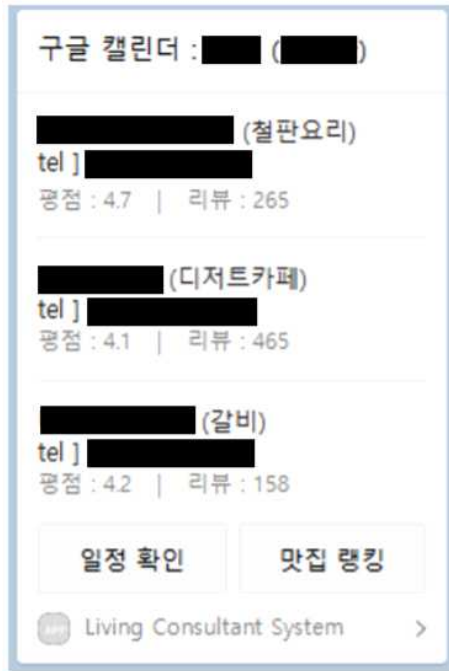
다음은 날짜, 시간에 대한 추출 방법이다. 텍스트에 한글로 표현된 정보가 있다면 이를 숫자로 변환시켜주는 과정을 거친다. '내일', '글피', '다음주' 등의 단어들이 이에 해당하게 된다. 그 이후엔 날짜와 시간을 판단하는 기준이 절대적인지, 상대적인지를 구분하여 구현하였다. 문장에서 '~뒤', '~후', '~부터', '~동안'이라는 워드가 언급되면 상대적인 시간 추출, 그 외의 경우에는 절대적인 시간을 추출하는 것으로 분류하였다. 분류 이후엔 표 2의 Rule에 따라 숫자에 대한 정보를 추출하였다.

6. 일정 등록 및 알림 자동화

텍스트에서 정보를 추출한 이후에는 이를 구글 캘린더에 등록하는 작업이 필요하다. 추출된 정보를 바탕으로 구글 캘린더 API를 통해 자동으로 일정을 등록해준다. 그 후엔 카카오톡 API를 이용하여 등록된 일정의 정보를 나에게 보내는 메시지로 전송하게 된다.



(그림 5) 일정 등록 직후 또는 장소가 등록되어 있지 않은 일정 알림 시 전송되는 메시지



(그림 6) 장소가 등록되어 있는 일정 알림 시 전송되는 메시지

아침 8시마다 당일의 일정에 대한 정보를 전송하도록 하였다. 구글 캘린더에서 당일의 일정 정보를 불러오고, 장소가 등록되어 있지 않은 경우에는 일정 등록 직후와 동일하게 일정명과 구글 캘린더 링크를 제공하였다. 장소가 등록되어 있는 경우에는 일정명과 장소, 구글 캘린더 링크, 맛집 사이트 링크를 제공하였다. 추가적으로 웹 크롤링 드라이버인 셀레니움[7]을 통해 카카오 맵에 장소+맛집을 검색한 후 크롤링하여 지정한 평점 이상의 식당만 선별하였다. 상표명, 카테고리, 연락처, 평점, 리뷰에 대한 정보를 바로 확인 가능하며, 링크로 연결되어 있어 자세한 정보도 확인 가능하도록 하였다.

처음 프로그램 실행 시, 구글 캘린더와 카카오톡 API를 사용하기 위한 인증 절차가 필요하다. 이를 간편화하기 위해 셀레니움으로 자동 토큰 발급 및 인증 자동화를 구현하였다.

7. 결론 및 향후 연구

본 연구에서는 음성 인식 기술을 통해 얻은 텍스트에서 정보를 추출하고, 자동 등록 및 알림, 추가 정보 제공해주는 시스템을 구축하였다. 핫워드를 통해 음성 인식을 시작하고, 음성으로 일정 등록에 대한 문장을 전달하면 자동으로 정보를 추출하여 구글 캘린더에 등록한다. 기존의 시스템을 사용하는 경우 날짜와 시간 외의 정보는 추출이 되지 않는 등의 문

제점이 있었으나, 본 연구에서 일정 등록에 필요한 정보를 기반으로 구조화하여 추출하는 것을 통해 해결하였다. 더불어 프로그램 사용 중에 생길 수 있는 인증에 대한 번거로움은 자동화를 통해 해결하였다. 본 연구에서 진행된 정보 추출 및 자동 일정 등록 시스템을, 추후에 홈 IoT 또는 어플리케이션을 연동한 정보 처리 등의 기능을 추가한 형태로 발전시킬 수 있다.

참고문헌

- [1] 김종훈, 송창우, 김주현, 정경용, 임기욱, 이정현, 음성인식을 이용한 상황정보 기반의 스마트 홈 개인화 서비스, 한국콘텐츠학회 논문지, 제 9권, 11호, pp.80-89, 2009
- [2] 남지순, 숫자포함 시간표현의 정보추출을 위한 언어 기술, 언어과학(JLS), 제 16권, 1호, pp.59-87, 2009
- [3] 최승주, 김종배, 음성 인식 오픈 API의 음성 인식 정확도 비교 분석, 예술인문사회 융합 멀티미디어 논문지, 제 7권, 8호, pp.411-418, 2017
- [4] <https://woongsin94.tistory.com/333>
- [5] 민소연, 이광형, 이동선, 류동엽, 한국어 특성 기반의 STT 엔진 정확도를 위한 정량적 평가방법 연구, 한국산학기술학회논문지, 제 21권, 7호, pp.699-707, 2020
- [6] 이정님, 음성 언어 처리 기술, 어디까지 왔나, 새국어생활, 제 27권, 4호, pp.99-116, 2017
- [7] <https://github.com/SeleniumHQ/selenium/releases>