

인공지능기반의 키워드 중심 회의록 자동 생성 시스템

*강수지, **유진주, ***이태림, ****이하연, *****임양미
 덕성여자대학교 IT미디어공학과
 astnwl0619@duksung.ac.kr

Automatic Creation of Artificial Intelligence Meeting Minutes System using Korean Keyword Extraction

*Kang Suji, **Yoo Jinjoo, ***Lee Taerim, ****Lee Hayeon, *****Lim Yangmi

Dept. of IT Media Engineering, Duksung Women's University

요 약

비대면 시대로 인한 화상 회의의 중요성이 높아졌다. 하지만 현재까지도 회의기록의 문서화 작업은 수작업으로 이루어지고 있어 시간과 인적자원이 많이 소모되고 있다. 본 논문은 기존 수작업으로 진행되는 회의 문서화 과정의 문제점을 개선하고자 한국어 키워드 추출을 활용한 인공지능 회의록 자동 생성 시스템을 제안한다. 회의 음성 파일을 기반으로 STT 기술을 활용한 회의 전문을 자동 생성하고 전문에 KR-WordRank 알고리즘을 적용해 키워드를 추출, Summary API를 사용하여 요약본을 생성한다. 최종 결과로 회의 전문과 키워드, 요약본이 담긴 PDF 형식의 회의록을 사용자에게 제공하여, 수기 회의록 작성 시 들이는 시간적, 인적 비용 절감을 돕는다.

1. 서론

비대면 업무가 증가함에 따라 화상 회의의 중요성이 높아졌다. 하지만 현재까지도 회의록 작성은 컴퓨터를 활용한 타이핑 기록 방식으로 작업되고 있다. 이에 소요되는 인적 노동력과 시간이 상당히 많이 소요되고 있는 상황이다.

본 논문은 장시간 회의 내용을 사람이 직접 요약하여 기록하던 회의록 타이핑 기록 방식을 자동화하기 위해 인공지능기반의 한국어 키워드 중심 회의록 자동화 시스템으로 개발하고자 한다.

Naver 기업이 개발한 음성기록 서비스 ‘클로바 노트(CLOVA)’는 음성 녹음을 텍스트화하여 사용자가 텍스트를 선택하면 원하는 부분의 음성 녹음을 다시 재생해 준다. 이러한 서비스는 대화 내용을 찾을 때 장시간 회의 음성 파일을 다시 들어야 하는 시간을 줄여주는 장점이 있다.[1] 구글의 Speech-to-Text API는 최첨단 자동 음성 인식(ASR) 딥 러닝 신경망 알고리즘을 적용하여 높은 정확성을 지원하고 화자 분할이 가능하다는 장점이 있지만 한국어에 적합하지 않아 네이버의 STT 기술을 본 연구에서 사용한다. 특히 CLOVA의 ‘클로바 스피치(CLOVA Speech)’ ‘네스트(NEST)’ 엔진은 대화의 긴 문장을 인식하는데 특화되어있어 장시간 회의 기록에 적합하다고 판단하였다.[3]

장시간의 회의 음성 파일에 STT를 활용하여 회의 전문을 텍스트 파일로 얻은 후, 전체 회의 내용을 요약하기 위한 키워드를 추출한다. 키워드 추출을 위해서 비지도학습 방식인 한국어 추출 알고리즘 KR-WordRank 을 활용하였다. 추출된 키워드를 중심으로 문장 요약이 되며 CLOVA Speech 의 Summary API를 사용하여 요약문을 자동 생성한다.

2. 작품의 설계 및 구현

인공지능 기반의 한국어 키워드 중심의 회의록 요약 자동화 시스템 개발을 위한 구성도는 fig 1과 같다.



Fig. 1 Service Diagram

제안한 시스템은 Python 기반 웹 프레임워크인 Django와 JavaScript 라이브러리인 React를 사용한다. 데이터베이스는 MySQL을 사용하여 CRUD를 구현 시 회의 정보를 저장하고 관리할 수 있도록 한다. 데이터베이스에 회의(Meeting) 모델, STT 결과(Result) 모델, 요약과 키워드(Summary) 모델을 생성한 후 조회, 추가, 수정, 삭제가 가능하다. 이후 세부적인 화면 구성과 UI/UX를 와이어프레임으로 설계한다.

Fig 1의 Naver Clova STT(speech to text) 단계에서는 음성 파일을 기반으로 한 회의 스크립트 제공한다. Naver CLOVA Speech API를 사용하여 구현한다. 키워드 추출 기능은 KR-WordRank 알고리즘을 적용한다.[2] 키워드 추출을 위해 전문의 전처리 과정을 거친다. Normalize 라이브러리를 사용해 불필요한 특수 기호들을 지우고 불용어 처리 사전을 통해 불용어를 제거한다. 단어의 최소 출현 빈도수와 단어의 최대 길이를 제한하여 그래프를 생성하고 그래프의 의미를 가지는

L-parts를 반환하여 키워드로 추출되도록 구현한다.

Fig 1의 구성도에서 가장 중요한 요약 기능은 CLOVA Summary API 및 KoBERT 모델 적용한 요약문 추출 학습 모델을 사용하여 요약문 생성이다. 생성된 회의록은 회의 정보, 요약, 키워드 모두 포함하여 제공하도록 설계하였고 PDF 형식의 회의록을 파일로 다운받을 수 있도록 구현한다.

3. 작품의 구현

1) Speech To Text(STT)

STT는 사람이 말하는 음성 언어를 컴퓨터가 해석해 그 내용을 문자 데이터로 전환하는 것을 말하며 음성 인식(Speech Recognition)이라고 한다. 1952년에 Bell 연구소에서 숫자 인식이 프로젝트를 시작으로 사칙연산 음성을 알아듣는 ‘Shoobox’, 일본에서 ‘모음’부분을 인식하는 하드웨어 개발, 한 단어가 어디에서 끝날지 판단하는 시스템 개발 등 점차 음성인식 가능 활용 범위가 넓어졌다. 현재 인간 속기사의 정확도 95% 정도의 성공률로 음성 인식 시장이 더 확대되었다.[1]

Naver Cloud Platform에서 제공하는 CLOVA Speech는 방송 영상이나 오디오 클립 등의 긴 음성 또한 높은 정확도를 통해 텍스트로 변환하는 기능을 제공하는 기술이다. 프로젝트에서 해당 기술을 사용하여 사용자로부터 받은 음성파일을 텍스트화하여 사용자에게 제공한다. 다음은 시간과 화자의 인원에 따른 결과물의 정확도에 대한 표이다.

[table 1] Accuracy of results

	length of audio(min.)	converting time(sec.)	number of characters	accuracy (%)
text 1	1	10	1,406	99.07
text 2	3	45	16,912	99.53
text 3	93	70	37,788	99.43

table 1의 text 1의 경우 말이 겹치는 부분이 없지만, 발화자가 말을 더듬으면서 정확도가 떨어졌다. text 2와 text 3의 경우 말이 겹치는 부분이 꽤 있지만, 발화자의 발음이 text 1에 비해 정확도가 높게 나왔다. Google Speech API의 경우 발화자가 많아지거나 음성이 긴 경우, 말이 겹치는 순간 정확도가 현저히 떨어진다. 그에 비해 CLOVA Speech API는 발음에 따라서만 정확성이 달라지는 것을 볼 수 있다.

2) Summary & KR-WordRank

요약문 제공에는 Naver Clova Summary API 모듈을 사용했다. 키워드 추출은 회의 전문 스크립트에 비지도 학습 기반 단어추출 알고리즘 모델 KR-WordRank를 사용했다. 다음 사진은 해당 API와 알고리즘 모델을 적용해본 예시이다.

SCRIPTS	SUMMARY
[여성]	김지에 변호사님께서 상당히 20대 여성의 어떤 취업의 힘든 고충에 대해서 깊이 공감을 많이 해 주심. 네, 그 이유는 이제 제가 만약에 할당제로서, 오직 성별만한 이유만, 성별만을 이유로 임원의 자리에 올라간다면 제가 그 자리에 있을 때 과연 제 발언권이 다른 남성분들과 동일하게 그분들께 설득력 있게 들릴 것인지에 대해서 먼저 궁금하고 또 추가로 아까 말씀해 주신 발언 중에서 이제 당장에, 20대를 당장의 여성들을 꾸어, 끌어줄 그런 임원이 부족하다는 말씀을 해 주심. 저는 저 저를 여자라고 끌어주길 바라지도 않고 만약 그렇게 끌어준다고 해서, 과연 현재로서 절대로 수치가 낮은데 그렇게 절대적으로 끌어준다고 해서 여성의 정상적인 양성평등의 향상이 과연 그쪽으로 이어질 수 있을지에 대해서 내 한번 여쭙보고 싶음.

Fig. 2 Result of System

핵심적인 스크립트 뽑아냈지만 첫째 화자가 부정확한 발화를 할 경우, 둘째 화자에 대한 답변자의 발화가 핵심 요약 결과물에 포함되지 않은 경우, 셋째로 음성의 길이가 짧아 키워드가 1개만 추출된 경우는 미흡한 결과를 보였다. (fig2. 참조).

table 2는 Naver Clova Summary API를 사용해 요약한 것과 TextRank로 같은 원문을 요약했을 때 비교할 수 있는 메트릭스인 ROUGE의 점수이다. 점수가 높지는 않지만 TextRank보다는 높은 점수를 내는 것을 확인할 수 있다.

[table 2] ROUGE's score

	Rouge-1	Rouge-2	Rouge-1
Summary API	0.45	0.41	0.45
TextRank	0.22	0.17	0.22

4. 결론

본 논문의 한국어 키워드 추출 기능을 활용한 인공지능 회의록 자동 생성 서비스로 회의록을 작성함으로써 수기 회의록 작성에 드는 시간과 노력을 대폭 줄일 수 있다. 대학생이나 소규모 모임뿐만 아니라 회의 빈도가 잦은 기업이나 공공기관에서 도입하여 더 큰 가치를 창출할 수 있는 잠재력을 갖고 있다. 회의뿐만 아니라 강의나 뉴스 등에 적용하여 텍스트로 변환할 수 있는 음성과 영상의 내용을 요약할 수 있기 때문에 소리를 들을 수 없는 청각 장애인들에게도 내용을 전달할 수 있을 것으로 기대 된다.

5. 참고문헌

- [1] K. Heo, J. Yang, D. Kim, K. Bok, J. Yoo, "Design and Implementation of Minutes Summary System Based on Word Frequency and Similarity Analysis", The Journal of the Korea Contents Association, Vol. 19, Issue 10, pp.620-629, Oct. 2019
- [2] H. Kim, S. Lee, "Korean Text Automatic Summarization using Semantically Expanded Sentence Similarity", Fall Conference 2014, Korea Information Processing Society, Nov. 05, pp.841 - 844, 2014, November
- [3] J. Lee, S. Lee, "Summary of meeting minutes using two-step sentence extraction", The Journal of the Korean Institute of Intelligent Systems, Vol. 20, Issue 6, pp.741 - 747, 2010