



음성데이터를 이용한 감정분석 프로그램 설계

Design of Emotion Analysis Program using Speech Data

저자 (Authors)	이상우, 윤영재, 이경희, 조정원 Sangwoo Lee, Yeongjae Yoon, KyungHee Lee, Jungwon Cho
출처 (Source)	한국정보과학회 학술발표논문집 , 2018.6, 2187-2188(2 pages)
발행처 (Publisher)	한국정보과학회 KOREA INFORMATION SCIENCE SOCIETY
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07503653
APA Style	이상우, 윤영재, 이경희, 조정원 (2018). 음성데이터를 이용한 감정분석 프로그램 설계. 한국정보과학회 학술발표논문집, 2187-2188
이용정보 (Accessed)	가천대학교 203.249.***.201 2019/09/29 19:00 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

음성데이터를 이용한 감정분석 프로그램 설계*

이상우⁰¹ · 윤영재¹ · 이경희² · 조정원¹

¹ 제주대학교 사범대학 컴퓨터교육과

² 제주대학교 대학원 과학교육학부 컴퓨터교육전공

tkddn1752@naver.com, scout125@naver.com, jwcho@jejunu.ac.kr, dreamer@jejunu.ac.kr

Design of Emotion Analysis Program using Speech Data

Sangwoo Lee¹ · Yeongjae Yoon¹ · KyungHee Lee² · Jungwon Cho¹

¹ Department of Computer Education, Jeju National University

² Major in Computer Education, Graduate School, Jeju National University

요 약

음성은 사람간의 의사소통에 있어 감정을 파악하기 좋은 방법 중 하나이다. 음성인식 감정분석 프로그램은 사람의 음성이 가진 특성을 이용하여 그 안에 담긴 감정을 분석한다. 음성이 가진 피치 값, 속도의 변화와 같은 요소를 데이터화 하여 그 안에 담긴 감정을 기계학습을 통해 분류 및 예측하는 과정을 거치게 된다. 감정 별 음성 데이터 분석을 위해 다양한 기계학습 알고리즘을 활용하여 기존 연구들보다 높은 정확도를 나타낼 수 있도록 연구, 개발 하였다.

1. 서 론

사람과 사람간의 의사소통에 있어 서로의 감정을 파악할 수 있는 수단은 목소리, 몸짓, 표정 등 다양한 방법이 존재한다. 이러한 방법 중 비교적 다른 수단들에 비해 목소리는 상대방의 감정을 파악하기 용이하다는 연구결과가 존재한다[1]. 음성분석을 통해 상대방의 감정을 유추할 수 있다는 결론을 도출할 수 있다. 기존 개발된 음성인식 감정분석 프로그램을 살펴보면 평균적으로 정확도가 60%밖에 되지 않을 정도로 낮은 편이다[2].

따라서 본 연구에서는 정확도 향상을 위해 기계학습 중 딥러닝(Deep-Learning) 기술을 사용해 정확도를 향상시키는 것을 목표로 한다. 또한 감정의 종류를 기쁨, 화남, 슬픔, 긴장, 놀람등 으로 분류하여 음성을 통한 감정분석에 효율적인 딥러닝을 진행할 수 있을 것으로 예상한다.

2. 이론적 배경

2.1 음성을 통한 감정 유추

인간의 음성을 통한 감정인식 분석 연구는 다양하게 존재해 왔다. 주로 음성신호에 들어있는 피치(PITCH)를 특정 정보로 활용한 감정 분석으로의 접근이 이뤄지고 있다[3]. 이를 기반으로 음성신호에서 감정을 반영하는 부분으로의 음성의 피치 값을 활용하고자 하였다.

대표적인 음성 기반 감정인식 연구의 적용사례는 콜센터의 모니터링 서비스이다. 이 연구는 통화가 이루어지는 동안 고객이 부정적인 감정을 표출할 경우 이를 감지하여 알려주는 역할을 한다. 이 연구는 통화 중 감정을 유추한다는 점에서 실시간 감정 인지를 제공한다.

이러한 음성을 기반으로 한 실시간 감정 인지 연구는 다양하게 존재한다. 그러나 기존 연구들은 특정 길이의 시간동안의 감정을 분석하거나 남/여와 같이 음성이 변인을 가지게 될 경우에 대한 고려를 필수로 한다[3].

한계점을 극복하기 위해, 기계학습 중 딥러닝을 기반으로 하여 감정 분석의 결과를 더욱 정확하게 하고자 하며, 다양한 음성 데이터를 활용한 감정 분석에 임하고자 한다.

* 이 논문은 2018학년도 제주대학교 지방대학 특성화사업(CK-1)에 의해 지원받았음

2.2 딥 러닝(deep-learning)

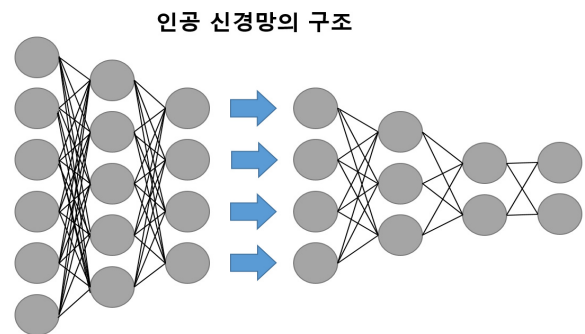


그림 1. 딥 러닝을 통한 유추과정

딥 러닝은 수많은 데이터를 분류해 같은 집합들끼리 묶고 상하의 관계를 파악하는 기술이다. [그림 1]은 딥 러닝의 인공신경망구조로 인간의 뉴런 구조를 본떠 만든 기계학습 모델이다. 데이터의 값들을 서로 비교분석 하면서 결과값을 예측하는 과정에 대한 설명이다.

딥 러닝의 핵심은 분류를 통한 예측이다. 수많은 데이터들 사이에서 패턴을 발견하여 그 패턴을 기반으로 결과 값을 예측해 주는 것이다. 딥 러닝에는 두 가지의 분별 방식이 있는데, 지도 학습과 비지도 학습으로 나뉜다. 지도 학습이란 정해진 정답을 미리 컴퓨터에게 알려주고 컴퓨터가 정답을 토대로 학습을 하는 방법이고, 비지도 학습은 데이터를 통한 학습을 통해 정답을 컴퓨터가 스스로 학습하는 방법이다. 본 연구에서는 데이터 셋이 갖추어지지 않은 상태이기 때문에 지도 학습을 선택하였다. 또한 딥 러닝은 사람의 신경망처럼 인공 신경망을 만들어 데이터의 결과 값을 예측하기 때문에 데이터의 양이 중요하다. [그림 1]처럼 데이터의 양이 많으면 많을수록 분별할 기준이 많아지기 때문에 정확도가 향상한다. 본 연구에서는 데이터를 직접 찾기보다는 이 프로그램 사용자들의 목소리를 다시 데이터화 시켜 딥 러닝 중에서 지도 학습을 할 것이다.

2.3 프라트(praat)

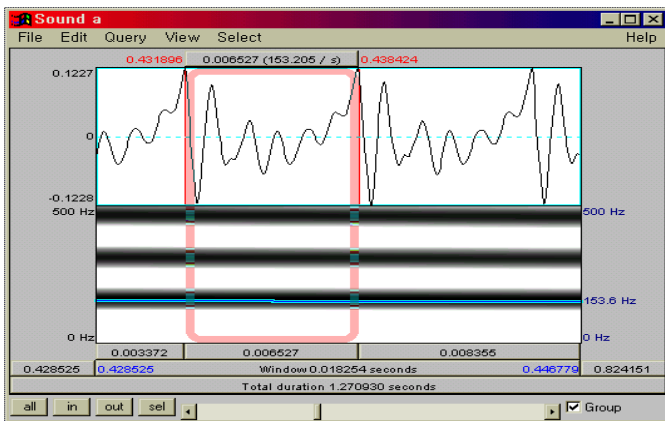


그림 2. 프라트의 피치 값 분석과정

[그림2] 는 프라트가 피치를 분석하는 과정 중 시간의 흐름에 따라 분석하는 과정이다[2].

프라트란 간단한 음성분석에서 복잡하고 전문적인 처리까지 가능한 도구로서 인터넷을 통해 무료로 다운로드 받을 수 있는 프로그램이다. 스크립트를 이용하여 대용량의 음성도 자동으로 처리할 수 있는 음성분석 프로그램이다. 프라트는 음성을 입력받으면 음성을 데이터화 시켜주고, 각 음성의 피치 값을 1/16초 단위로 추출해준다. 프라트를 이용해 나온 피치 값을 기준으로 감정분석에 사용하려고 한다[4].

3. 딥 러닝을 활용한 음성 기반 감정 분석 프로그램 설계

3.1 음성 기반 감정 분석 단계

음성을 이용해 감정을 분석하기 위해, 음성신호에서 피치(PITCH) 값을 추출하도록 한다. 이 과정은 프라트(Praat) 프로그램을 활용할 것이며 이를 통해 얻어낸 음성 데이터 값을 기계학습을 이용하여 분류할 수 있도록 한다.

음성 데이터를 수집하여 파이선을 이용한 기계학습을 진행한다. 이 때, 기계학습 중 딥 러닝의 특징을 활용, 데이터가 많을수록 높은 정확도를 나타내줄 수 있도록 한다.

3.2 프라트(Praat)를 통한 음성 데이터 분석

음성을 이용하여 감정을 분석하기에 앞서, 음성 신호를 데이터로 변환해 줄 필요가 있다. 이 과정은 프라트 프로그램을 통해 이루어질 것이며, 음성 신호를 토대로, 감정 분석에 필요한 피치 값을 추출해내는 것이다.

[그림 3]은 프라트를 통해 음성 데이터를 분석한 결과이다.

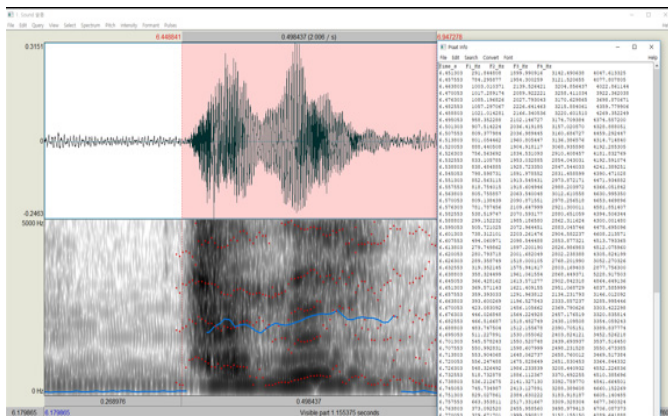


그림 3. 프라트의 피치 값 분석결과

3.3 기계학습(머신러닝)을 통한 감정 예측

파이선의 각종 기계학습 라이브러리를 활용하고자, 아나콘다 3와 eclipse를 개발환경으로 선정하였다. 프라트를 통해 변환한 음성 데이터를 이용하여 딥 러닝을 진행한다. 해당 데이터들은 각자 피치 값의 변화를 비롯한 음성 신호들을 이용하여 기쁨, 슬픔, 화남, 놀람 등의 감정별 음성이 가지는 특성으로 분류하여 일정량의 데이터를 입력하게 된다. 차후 기계학습을 통해 새롭게 입력받은 음성 신호를 각 감정별 특성에 따라 분류, 예측하여 결과를 나타내게 된다.

이 때 활용하게 될 기계학습 라이브러리는 인공지능 분야의 발전을 위해 구글에서 내놓은 공개 라이브러리인 ‘텐서 플로우(Tensor flow)’를 비롯하여 mglearn과 같은 각종 파이선의 라이브러리들을 활용하도록 한다.

4. 결론 및 향후 연구

음성이 갖게 되는 특성을 분류하여 이를 토대로 기계학습을 진행하여 보다 높은 정확도의 감정 예측 결과를 유추할 것이다. 높은 정확도를 위해 더 많은 양의 데이터 수집을 목표로 하여, 다양한 사용자들의 음성 데이터를 추가적으로 활용하는 방안을 고려한다. 또한, 기존에 활용하는 기계학습 알고리즘 외에도 추가적인 기계학습 관련 연구를 진행하도록 한다.

향후 연구에서는 본 연구에서 제작된 프로그램을 애플의 시리나, 삼성의 빅스비 같은 지능형 개인 비서 소프트웨어에 적용하여, 단순한 대답이 아닌 이용자의 감정까지 분석해 대답을 하는 연구를 할 수 있을 것이다. 또한 본 연구의 감정 중 하나인 긴장의 감정을 이용해, 보이스피싱과 같은 범죄 예방에 적용할 수 있을 것으로 기대한다.

5. 참고 문헌

- [1] 마이클클라우스(Michael Kraus), 미국심리학회(American Psychologist Association), 어메리칸 사이콜로지스트 저널(American Psychologist) 2017.
- [2] 방재훈, 이승룡. 감성기반 서비스를 위한 통화 음성 감정 인식 기법, 정보과학학회논문지 : 소프트웨어 및 응용, 제 41권 제 3호, 2014.
- [3] 김남균, 성우경, 하현규, 김홍국, 음성신호 기반 감정인식을 위한 특징 파라미터 비교, 한국통신학회, 2016.
- [4] 양병곤, 프라트(Praat)를 이용한 음성분석의 이론과 실제, 만수출판사, 2010.