

멀티모달을 적용한 하이브리드 감정인식

임명진¹, 신주현^{2*}

조선대학교 컴퓨터공학과¹, 조선대학교 신산업융합학부^{2*}

e-mail : dolly720@hanmail.net, jhshinkr@chosun.ac.kr

Hybrid Emotion Recognition using Multi-Modal

Myungjin Lim¹, JuHyun Shin^{2*}

Dept. of Computer Engineering, Chosun University¹,

Dept. of New Industry Convergence, Chosun University^{2*}

요 약

최근 COVID-19로 인한 비대면 서비스의 증가로 SNS로 소통하는 경우가 점점 늘어나고 있다. SNS 데이터는 사용자 정보나 의견을 분석하고 감정을 인식하는 연구로 많이 활용되고 있다. 현재 텍스트 감정인식은 대부분 단어나 문장 단위의 단일 감정으로 분류되고 있다. 그러나 하나의 문장에도 여러 감정이 복합적으로 존재하고 연속적인 문장에는 감정의 흐름이 존재하기 때문에 더 정확하게 감정을 인식하는 방법이 필요하다. 따라서 본 논문에서는 더욱 정확한 텍스트 감정인식을 위해 키스트로크 다이나믹스를 활용하여 멀티모달을 적용하고 다중 감정과 연속 감정을 결합한 하이브리드 감정인식 방법을 제안한다.

1. 서 론

최근 COVID-19로 인한 비대면 서비스의 증가로 SNS를 통해 자신의 의견을 표현하고 경험을 공유하며 소통하는 경우가 점점 늘어나고 있다. 사용자에게 의한 대량의 SNS 데이터가 발생하면서 사용자 정보나 의견을 분석하고 감정을 인식하는 연구가 활발하게 진행되고 있다[1]. 그중에서도 감정인식은 사회적 이슈나 제품에 대한 개인적인 의견을 분석하여 감정을 인식하는 연구 분야이다. SNS에서 텍스트 데이터는 특성상 보내는 사람과 받는 사람 간에 다른 의미로 해석되어 뜻하지 않게 오해가 생길 수 있고, 보낸 사람이 의도한 문맥적인 상황이 받는 사람에게는 제대로 전달되지 않을 때도 있다. 대화 당사자 간 오해가 없는 대화를 위해서 투명한 감정 교류는 중요한 요소이다. 현재 텍스트 감정인식은 대부분 단어나 문장 단위의 단일 감정으로 분석되고 있다. 그러나 하나의 문장에도 여러 감정이 복합적으로 존재하고 연속적인 문장에는 감정의 흐름이 존재하기 때문에 더 세밀하게 분석하여 정확한 감정을 인식하는 방법이 필요하다. 따라서 본 논문에서는 더욱 정확한 텍스트 감정인식을 위해 키스트로크 다이나믹스 기술을 활용하여 멀티모달을 적용하고 문장의 다중 감정과 연속적인 문장의 연속 감정을 결합한 하이브리드 감정인식 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구로 텍스트 기반 감정인식과 멀티모달에 관해 기술하고, 3장에서는 본 논문에서 제안하는 멀티모달을 적용한 하이브리드 감정인식 방법에 관해 기술한다. 마지막으로 4장에서는 결론과 향후 연구에 관해 기술하고 마무리한다.

2. 관련 연구

감정인식은 모달리티의 종류에 따라 음성, 생체신호, 비전, 텍스트 등으로 분류할 수 있다. 다른 모달리티보다 텍스트 기반 감정인식은 단순히 현재 입력 데이터에 대해 감정 분류를 해주는 수준을 넘어서 과거 기억이나 감정 주제, 성격이나 성향 등 지능적으로 더욱 정확한 감정인식을 가능하게 하는 데 필요하다[2]. 과거에는 대부분 감정 키워드를 추출하거나 어휘를 통해 감정을 분류하였으나 현재는 주로 딥러닝을 사용하여 문장의 감정을 분류한다[3,4]. 하지만 대부분 단어나 문장 단위의 단일 감정으로 분류된다는 한계가 있다.

키스트로크 다이나믹스는 사용자가 어떤 키를 언제 눌렀고 언제 뗐는지에 대한 시간 정보를 다룬다. 하나의 키를 눌렀다가 뗄 때 걸린 시간, 키를 누르고 다음 키를 누를 때까지의 시간, 문장 입력 시 걸린 총시간, 키 사용 빈도 등을 측정할 수 있다[5]. 사용자가 입력한 텍스트와 키스트로크 다이나믹스를 결합한 멀티모달을 적용하면 더욱 정확한 감정인식이 가능하다.

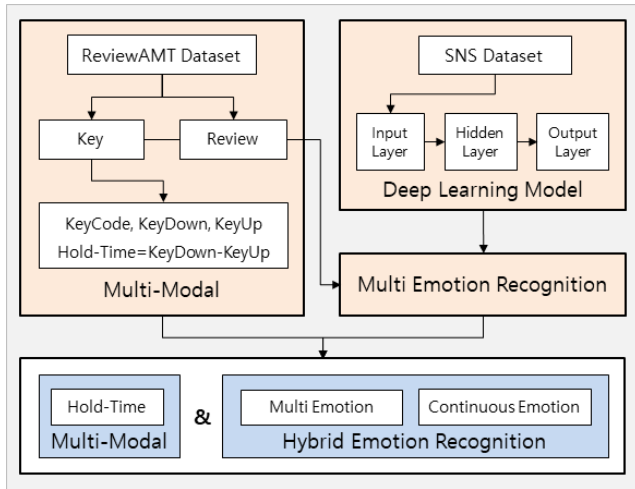
본 논문에서 제안하는 방법은 멀티모달을 적용하는 방법과 딥러닝을 사용하여 문장의 다중 감정을 인식하는 방법, 연속적인 문장에서 감정의 흐름을 인식하는 방법을 적용하여 더욱 정확한 감정인식을 가능하게 하는 것이다.

3. 멀티모달을 적용한 하이브리드 감정인식

3.1 시스템 구성도

본 장에서는 감정인식의 정확성 향상을 위해 키스트로크 다이나믹스를 활용하여 멀티모달을 적용하고 문장의 다중 감정과 연속적인 문장의 연속 감정을 결합한 하이브리드

리드 감정인식 방법을 제안한다. 그림 1은 본 논문에서 제안하는 방법의 시스템 구성도이다.



(그림 1) 시스템 구성도

그림 1과 같이 ReviewAMT 데이터 셋에서 Key와 Review 데이터를 구분하고 Key 데이터에서 하나의 키를 눌렀다 놓는 시간인 Hold-Time을 추출하여 멀티모달로 사용한다. 감정이 분류되어있는 SNS 데이터 셋을 딥러닝 학습하여 다중 감정인식 모델을 만들고 Review 데이터의 다중 감정을 예측한다. 키스트로크 다이내믹스를 적용한 멀티모달과 문장의 다중 감정, 연속적인 문장의 감정의 흐름을 결합한 하이브리드 감정인식 방법을 활용하여 더욱 정확한 감정을 인식한다.

3.2 데이터 셋과 전처리

실험을 위한 데이터는 ReviewAMT 데이터 셋[6]을 사용하였다. 미국의 레스토랑 리뷰로 텍스트와 키스트로크 다이내믹스 데이터로 구성되어 있으며 총 2,363개의 리뷰로 구성되어 있다. 전처리 단계로 텍스트는 Review 열로, 키스트로크 다이내믹스는 Key 열로 구분한 후 Key 데이터에서 KeyCode에 따라 KeyDown(키를 누르는 시간)과 KeyUp(키를 놓는 시간)의 차이를 계산하여 Hold-Time을 추출한다. Review와 KeyCode, KeyDown, KeyUp, HoldTime 열로 구분하였고 이 데이터 셋을 활용하여 KeyCode에 따른 Hold-Time을 분석할 수 있다. 그림 2는 전처리 후 데이터 셋을 나타낸다.

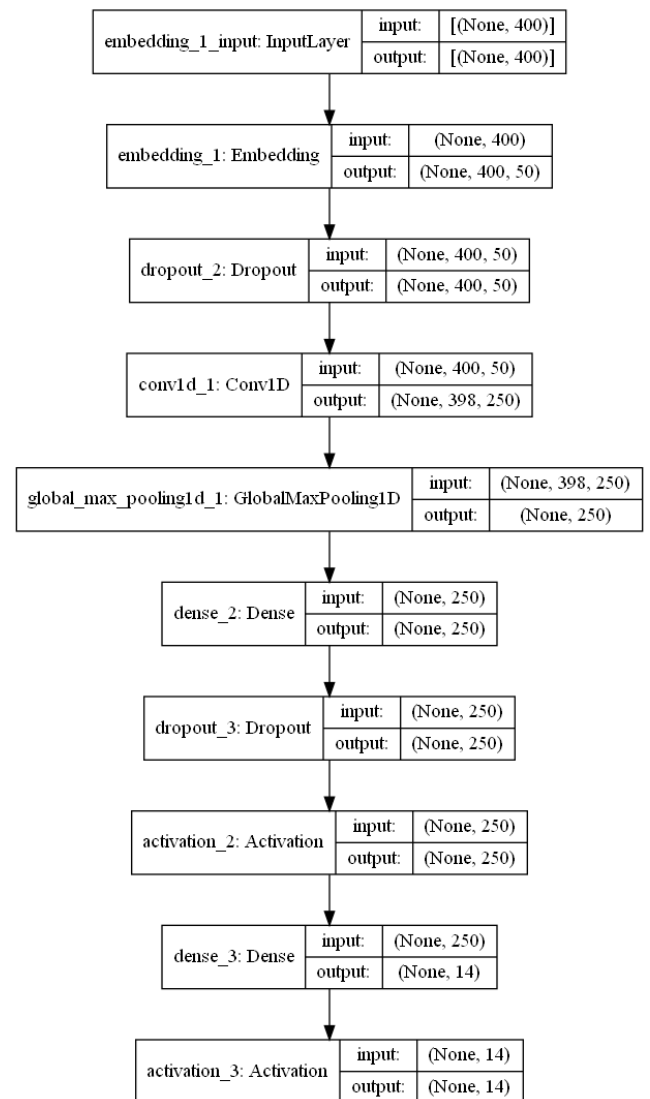
	A	B	C	D	E
1	Review	KeyCode	KeyDown	KeyUp	HoldTime
2	Famous Daves is a good	16	535	795	260
3	place to go for some	70	776	968	192
4	good quality food. They	65	924	1051	127
5	have an assortment of	77	1025	1171	146
6	barbecued food such as	79	1266	1405	139
7	ribs chicken and steak.	85	1369	1463	94
8	They also have a large	83	1505	1613	108
9	assortment of various	32	1909	2015	106
10	side dishes. The staff at	16	2065	2235	170

(그림 2) 전처리 후 데이터 셋

3.3 멀티모달과 다중 감정인식

전처리한 데이터 셋에서 하나의 Review는 여러 개의 연속 문장으로 구성되어 있고 그중 Review 1은 8개의 문장으로 구성되어 있다. Review를 문장으로 분류하고 마침표에 해당하는 KeyCode인 140을 기준으로 Hold-Time도 분류한다. 그리고 문장의 Hold-Time 평균을 계산하여 멀티모달로 적용한다.

ReviewAMT 데이터 셋은 Review와 Key로 구성되어 있어서 Key 데이터를 활용하여 멀티모달을 적용할 수 있지만, Review에 따른 감정은 분류되어있지 않는다. 따라서 본 논문에서는 감정이 분류되어있는 SNS 데이터 셋 [7]을 딥러닝 학습하여 감정인식 모델을 만든 후 이 모델을 활용하여 Review 데이터의 다중 감정을 예측한다. 14개의 감정(anger, boredom, empty, enthusiasm, fear, fun, happiness, hate, love, neutral, relief, sadness, surprise, worry)으로 분류된 40,000개의 SNS 데이터 셋을 딥러닝 학습하여 감정인식 모델을 만든다. 그림 3은 다중 감정인식 딥러닝 모델을 나타낸다.



(그림 3) 다중 감정인식 딥러닝 모델

그림 3과 같이 SNS 데이터 셋은 Embedding layer를 통해 400차원으로 임베딩 한 후 딥러닝 모델을 학습한다. 학습한 딥러닝 모델은 250개의 필터를 가진 Convolution layer와 MaxPooling layer를 통해 데이터의 특징을 추출하고 과적합 방지를 위해 Dropout을 사용한다. 출력층에 사용된 MLP는 다중 분류를 하기 위해 감정의 수인 14개로 출력하고 활성화 함수는 softmax를 사용하여 감정 분류를 진행한다. 학습 결과 감정 분류 정확도는 87%로 높은 정확도를 보여준다. 다중 감정인식 딥러닝 모델을 활용하여 Review 데이터 셋에서 문장별 다중 감정을 예측한 결과 14개의 감정으로 분류가 되었다. 표 1은 Review 1의 문장별 다중 감정과 Hold-Time을 나타낸다.

(표 1) Review 1의 문장별 다중 감정과 Hold-Time

No	Multi Emotion				Hold-Time
	Large1		Large2	...	
1	enthusiasm	78.68	happiness	19.91	123.88
2	sadness	84.73	hate	6.69	96.17
3	neutral	97.54	worry	2.45	101.43
4	fun	79.28	worry	19.57	111.63
5	worry	55.68	sadness	20.72	84.10
6	worry	99.01	surprise	0.07	95.65
7	worry	80.09	sadness	15.50	91.13
8	fun	86.64	fear	12.58	112.00

Review 1의 문장별 다중 감정 중 가장 큰 감정(Large1)과 두 번째로 큰 감정(Large2), 그리고 Hold-Time을 나타낸다. 이를 분석해보면 1번 문장에서 가장 큰 감정은

enthusiasm(열정)으로 78.68%이고, 두 번째로 큰 감정은 happiness(행복)로 19.91%를 차지하며, Hold-Time의 평균은 123.88인 것을 알 수 있다. 표 2의 결과를 활용하면 문장에 내제되어있는 여러 가지 다중 감정과 Hold-Time과의 연관성을 분석할 수 있다.

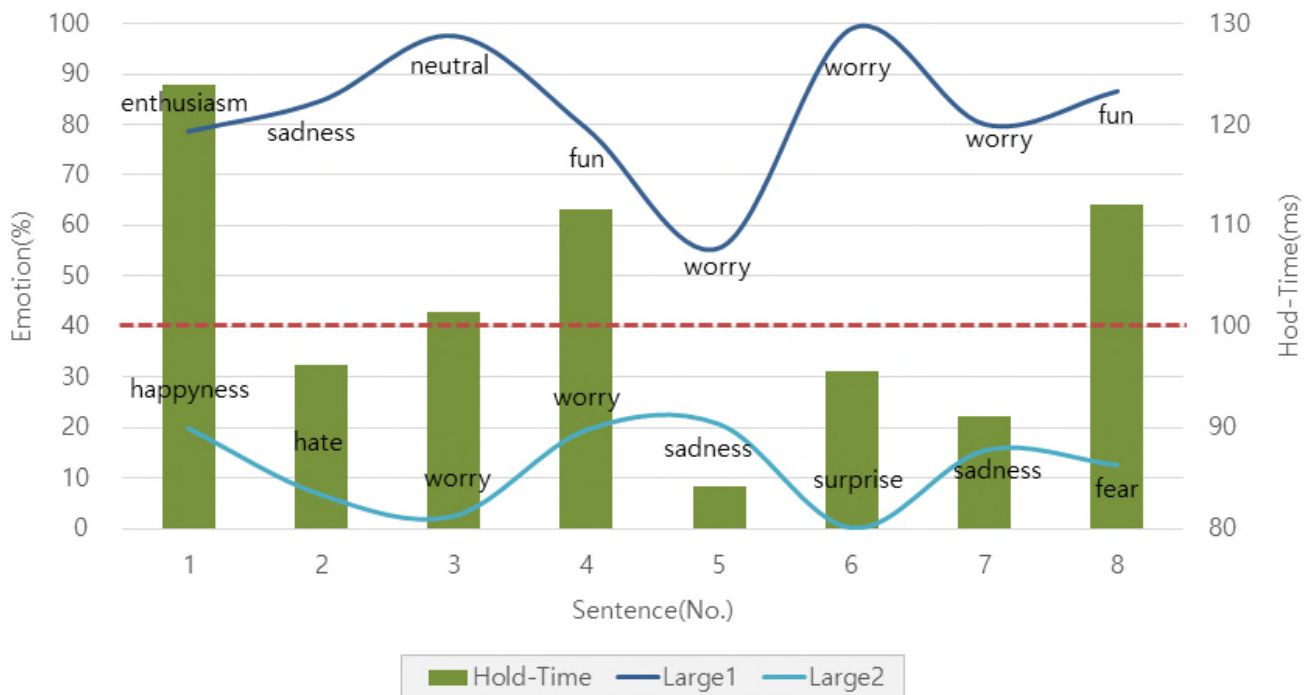
3.4 하이브리드 감정인식

앞에서 분석한 문장의 다중 감정과 Hold-Time을 활용하여 하이브리드 감정을 인식한다. 문장의 다중 감정을 이용하면 연속적인 문장에서 감정의 변화인 연속 감정도 인식할 수 있다. 그리고 다중 감정과 연속 감정을 결합하면 하이브리드 감정인식이 가능하다.

그림 4는 표 2의 분석 결과를 그래프로 나타낸 것이다. 문장별 Hold-Time과 가장 큰 감정(Large1)의 변화, 두 번째로 큰 감정(Large2)의 변화를 표시하였다. Review 1의 문장별 Hold-Time 평균과 하이브리드 감정을 분석해보면 Hold-Time 100을 기준으로 1, 4, 8번과 같이 긍정의 감정이 높을 때는 그 이상이고 2, 5, 6, 7번과 같이 부정의 감정이 높을 때는 그 이하이며 3번과 같이 중립의 감정이 높을 때는 100에 가까운 것을 볼 수 있다. 또한 문장에 내제된 다중 감정을 통해 연속적인 문장에서 감정의 변화를 볼 수 있다. 따라서 멀티모달을 적용한 하이브리드 감정인식 방법을 활용하면 더욱 정확한 감정인식이 가능하다.

4. 결론 및 향후 연구

본 논문에서는 텍스트 감정인식의 정확성 향상을 위해 멀티모달을 적용한 하이브리드 감정인식 방법을 제안한다. ReviewAMT 데이터 셋에서 Review와 Key 데이터를 구



(그림 4) Review 1의 Hold-Time과 하이브리드 감정

분하고 Key 데이터에서 Hold-Time을 추출한다. 감정이 분류된 SNS 데이터 셋을 딥러닝 학습하여 다중 감정인식 모델을 만들고 Review 데이터의 다중 감정을 예측한다. Hold-Time을 이용하여 멀티모달을 적용하고 문장의 다중 감정과 연속 감정을 결합하여 하이브리드 감정을 인식한다. 감정인식 결과 문장별 Hold-Time은 긍정일 때는 100 이상이고 부정일 때는 100 이하인 것을 볼 수 있다. 그리고 문장에 내제되어있는 다중 감정을 인식하면 연속적인 문장에서 감정의 흐름을 더욱 정확하게 파악할 수 있다. 향후 연구로는 연속 감정을 학습하여 딥러닝 모델을 만들고 예측하는 시스템을 구축할 계획이다.

감사의 글

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1F1A1057325).

참고 문헌

- [1] Shuai Li, Fei Hao, Hee-Cheol Kim. "Online Social Media Review Mining for Living Items with Probabilistic Approach: A Case Study," 『스마트미디어 저널』 제 2권 제 2호, 2013. pp.20-27.
- [2] 설용수, 김동주, 김한우, 박정기. "Predicate Logic Form을 이용한 자연어 텍스트로부터의 감정인식", 『한국 컴퓨터정보학회 학술발표논문집』 제 18권 제 2호, 2010. pp.411-412.
- [3] Matthew Purver and Stuart Battersby, "Experimenting with Distant Supervision for Emotion Classification", In Proceedings of EACL, 2012.
- [4] 신동원. "CNN-LSTM 복합 모델을 이용한 대화의 사용자 발화 감정 분류", 석사, 고려대학교, 2017.
- [5] 서동민, 강필성. "자유로운 문자열의 키스트로크 다이 나믹스와 일범주 분류기를 활용한 사용자 인증", 『대한산 업공학회지』 제 42권 제 4호, 2016. pp.280-289.
- [6] "Keystroke Patterns as Prosody in Digital Writings: A Case Study with Deceptive Reviews and Essays", Keystroke Patterns, 2020년 7월 5일 접속, <https://www3.cs.stonybrook.edu/~rbanerjee/project-pages/keystrokes/keystrokes.html>.
- [7] "mankutimma/text-emotion-classification", GitHub, 2021년 10월 1일 접속, <https://github.com/mankutimma/text-emotion-classification>