

## 딥러닝 기반의 다범주 감성분석 모델 개발\*

알렉스 샤이코니\*\* · 서상현\*\*\* · 권영식\*\*

### Development of Deep Learning Models for Multi-class Sentiment Analysis\*

M. Alex Syaekhoni\*\* · Sang Hyun Seo\*\*\* · Young S. Kwon\*\*

#### ■ Abstract ■

Sentiment analysis is the process of determining whether a piece of document, text or conversation is positive, negative, neutral or other emotion. Sentiment analysis has been applied for several real-world applications, such as chatbot. In the last five years, the practical use of the chatbot has been prevailing in many field of industry. In the chatbot applications, to recognize the user emotion, sentiment analysis must be performed in advance in order to understand the intent of speakers. The specific emotion is more than describing positive or negative sentences. In light of this context, we propose deep learning models for conducting multi-class sentiment analysis for identifying speaker's emotion which is categorized to be joy, fear, guilt, sad, shame, disgust, and anger.

Thus, we develop convolutional neural network (CNN), long short term memory (LSTM), and multi-layer neural network models, as deep neural networks models, for detecting emotion in a sentence. In addition, word embedding process was also applied in our research. In our experiments, we have found that long short term memory (LSTM) model performs best compared to convolutional neural networks and multi-layer neural networks. Moreover, we also show the practical applicability of the deep learning models to the sentiment analysis for chatbot.

Keyword : Sentiment Analysis, Convolutional Neural Networks, Long Short-Term Memory, Word2vec

## 1. 서론

감성분석(sentiment analysis)은 텍스트 데이터가 표상하는 작성자의 의견이나 감성, 평가, 태도 등을 분석하는 일련의 과정을 의미한다(Liu, 2012). 특히, 소셜미디어의 발전으로 인한 텍스트 데이터에 대한 의미변화는 감성분석의 필요성을 더욱 높이고 있다. 비정형 텍스트 데이터는 분석이 어려울 뿐 아니라 분석의 결과로 얻을 수 있는 가치가 한정적이라고 인식되었던 것에 반해 최근에 소셜미디어 등의 발전으로 인해 텍스트 데이터는 개인의 취향이나 성향 등을 파악할 수 있는 중요한 데이터로 인식되는 것에서 기인한다. 즉, 감성분석을 통해서 소셜미디어의 게시물, 인터넷 뉴스의 댓글, 상품 리뷰 등을 분석함으로써 특정 대상에 대한 텍스트 작성자의 의견을 손쉽게 파악할 수 있다. 이러한 이유로 대규모의 텍스트 데이터를 자동으로 신속하고 정확하게 분석하기 위한 감성분석 기법에 대한 연구의 필요성이 증대되고 있다.

소셜 미디어 이외에도 다양한 분야에서 감성분석을 활용하고 있으며, 관련연구 또한 지속적으로 이루어지고 있다(Feldman, 2013). 사회기사 및 상품 리뷰 등에서 감성분석이 활용되고 있으며, 이를 통해 비즈니스에 대한 인사이트를 제공하는 한편 상품 등에 대해서 고객의 즉각적인 피드백을 얻을 수 있다(Pang and Lee, 2008). 이외에도 감성분석은 정치 캠페인 등에 이용이 가능하며, 특정 정책이나 후보의 선호도 등을 파악하는 하나의 방법으로 대두되고 있다. 실제로 트위터 등의 소셜미디어의 텍스트 데이터에 대해서 정당과의 관계를 분석한 연구가 진행 된 바 있다(Bakliwal et al., 2013). 이외에도 감성분석을 이용하여 추가예측을 시도하는 연구가 진행되는 등 다방면에서 감성분석의 활용이 이루어지고 있다(Bollen et al., 2011). 최근에는 이용자와 자동으로 대화를 이어가는 챗봇 등의 연구가 활발히 진행되며 대화 시스템에서의 감성분석에 관한 연구 또한 시도되고

있다.

이와 같이 감성분석은 개개인의 일반론적인 감성을 파악하는 것을 넘어서 사회의 각 분야에서 활용이 가능하다. 분석 대상이 되는 텍스트 데이터의 범위 또한 문서, 문단 등의 상대적으로 길이가 긴 텍스트데이터부터 소셜미디어의 짧은 단문에 이르기 까지 다양하다. 일반적으로 감성분석은 대상 텍스트 데이터에 대한 감정을 긍정과 부정 등의 이진으로 분류를 수행한다. 일반적인 감성분석 방법론이 단순히 이진 분류를 수행할지라도 텍스트 데이터의 정보가 많은 문서나 문단을 이용하여 하나의 감성으로 요약하는 것은 쉽지 않다. 상대적으로 길이가 짧은 단문의 대화로 이루어진 챗봇 등에서는 반대로 적은 정보에서 감정을 요약하기 때문에 제약이 발생한다. 또한 감성분석을 효율적으로 사회분야에 응용하기 위해서는 긍정과 부정의 이진분류를 넘어 다양한 감성을 분석할 수 있는 다범주 감성분석 방법론이 필요하다.

이러한 상황에서 감성분석의 효율을 높이기 위해 딥러닝모델을 감성분석에 이용하는 연구들이 시도되고 있다. 특히, 합성곱신경망(convolutional neural networks : CNN), 순환신경망(recurrent neural networks) 등의 딥러닝모델들이 감성분석에 적용되어 기존의 감성분석 방법론들의 성능을 상회하는 결과를 보이고 있다. 하지만 딥러닝모델에 의한 감성분석 연구의 대부분은 긍정과 부정의 이진분류를 중심으로 연구가 진행되고 있어 다범주 감성분석에 대한 연구는 거의 없는 상황이다.

이에 본 연구에서는 딥러닝을 이용한 다범주 감성분석 모델을 제안한다. 제안하는 모델은 다범주 레이블을 가진 텍스트 데이터에 대해서 word2vec 신경망 기반 워드임베딩 모델을 통해서 워드임베딩을 수행하여 텍스트 데이터를 특정 벡터공간의 벡터로 임베딩 하고, 벡터로 표현된 텍스트 데이터를 이용하여 합성곱신경망(CNN), Long short term memory(LSTM) 모델을 학습시켜 텍스트 데이터를 분류하고자 하였다

## 2. 관련 연구

### 2.1 감성분석

기존의 감성분석은 크게 세 가지 접근방식으로 수행되어왔다. 첫째는 감성어 사전을 이용하여 감성분석을 수행하는 어휘기반 접근방식이고(Ding et al., 2008; Taboada et al., 2011), 둘째는 기계학습 기법들을 적용하여 감성을 분류하는 기계학습 기반 접근방식이며(Tan et al., 2009; Mullen et al., 2004), 세 번째는 어휘기반 접근방식과 기계학습 기반의 접근방식을 함께 사용하는 혼합 접근방식이다(Medhat et al., 2014).

어휘기반 접근방식은 단어 단위로 이루어진 문장이나 문단, 문서의 감성분석을 수행할 경우, 그 기준을 미리 사전으로 정의한 어휘에 두는 기법을 의미한다. 따라서 어휘기반 접근방식으로 감성분석을 수행하기 위해서는 미리 감성사전을 정의한 뒤, 감성분석의 대상이 되는 텍스트 데이터의 감성 점수를 계산하여, 긍정 혹은 부정으로 분류하게 된다. 이 방법은 감성사전이 구축된 이후에는 비교적 쉽게 감성분석을 수행할 수 있으나, 미리 감성어 사전을 구축해야만 하며, 사전에 등재되지 않은 새로운 감성어를 분석해야 할 경우에는 분류가 어려운 단점이 있다. 또한 단어 등의 어휘에 기반을 두어 감성분석을 수행하기 때문에 어휘보다 상위인 문장, 문단, 문서 등의 의미를 온전히 반영하기 어려운 단점이 있다.

기계학습 기반의 접근방식은 텍스트 데이터에 대해서 기계학습 방법론을 활용할 수 있도록 특징을 추출하고, 추출한 텍스트 데이터의 특징들을 바탕으로 기계학습 모델을 학습시켜 감성분석을 수행하는 모델을 의미한다. 기계학습 기반의 감성분석은 감독(supervised) · 비감독(unsupervised) · 반감독(semi-supervised) 학습의 다양한 기법들을 통해 수행된다. 그 중에서도 감성분석에 활용되는 감독 학습의 대표적인 기법으로는 나이브 베이저안 분류기(naive bayesian classifier), 서포트 벡터 머신

(support vector machine), 의사결정나무(decision tree), 인공신경망(artificial neural networks) 등이 있다(Pang et al., 2002).

최근에는 기계학습 접근방식의 하나의 방법으로서 딥러닝 기반의 심층신경망을 이용한 감성분석에 대한 연구가 수행되고 있다. 특히, 합성곱신경망 등을 통한 긍정과 부정의 이진분류 감성분석이 수행되어 의미 있는 성과를 거두기도 하였다(Kim, 2014). 또한, 순환신경망의 장기기억문제를 셀구조를 통해 효과적으로 해결한 LSTM 모델을 활용하여, 감성분석이 시도되기도 하였다(Tai et al., 2015). 이와 같은 딥러닝 모델들이 제안하였지만, 이들 연구는 이진분류를 위한 모델이다.

감성분석 모델은 학습데이터와 실제 입력되는 데이터 간의 차이가 있을 경우에는 성능이 하락하는 단점이 있다. 이와 같은 문제의식을 바탕으로 기존의 감성분석 방법론에서 더 나아가 특정 텍스트 데이터를 다양한 레이블로 분류하고자 하는 시도들이 있으나(Liu et al., 2017), 아직까지 감성분석과 관련된 주요 연구는 이진분류 모델을 중심으로 이루어지고 있기 때문에 실제 응용의 실용성 측면에서 한계가 있다. 이와 같은 이유로 다범주 감성분석 모델에 대한 활용성과 실용성이 증대되고 있는 실정이다.

### 2.2 딥러닝 기반의 감성분석 방법론

텍스트 데이터를 분석하기 위해서는 문자로 구성된 데이터를 숫자의 형태로 변환해서 표현해야 한다. 워드 임베딩(word embedding)은 텍스트 데이터를 숫자로 구성된 특정한 벡터 공간으로 사상시키는 과정을 의미한다. 워드 임베딩을 통한 단어의 표현방식은 국소표현(one-hot representation)과 분산표현(distributed representation)이 있다. 국소표현은 n개의 단어집합이 있을 때, 특정 차원만 1로 나머지를 0으로 표현하여 n차원의 벡터로 표현하는 방식을 의미한다. 최근에는 신경망 기반의 워드임베딩 기법인 word2vec이 제안되어 주목 받아 있다(Mikolov et al., 2013). word2vec은 문장

내에서 목표단어 전후에 배치되어 있는 단어들을 바탕으로 목표단어를 예측하는 CBOW(continuous bag-of-words)와 특정 단어를 입력으로 하여 그 단어 전후의 단어들을 예측하는 Skip-gram(continuous skip-gram)의 두 모델을 일컫는다.

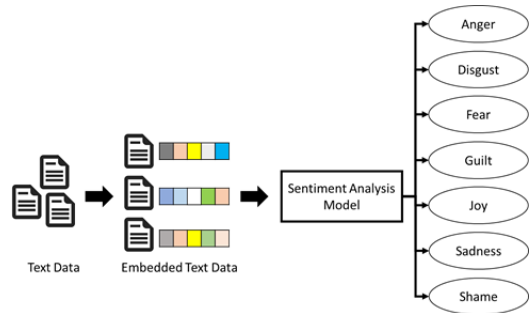
합성곱신경망은 AlexNet 등의 모델 등이 제안된 이후로 이미지처리와 관련된 분야에서 주로 사용되는 딥러닝모델의 하나로서 많은 주목을 받고 있는 신경망 구조이다(Krizhevsky et al., 2012). 합성곱신경망은 감성분석의 영역에도 적용되어 긍정과 부정의 이진 분류의 영역에서 의미 있는 성과를 거두고 있다(Kim, 2014; Dos Santos et al., 2014; Poria et al., 2015).

순환신경망은 일반적으로 입력문장이 길어질 경우 성능이 저하된다. 긴 문장에 대한 순환 신경망의 성능 저하는 신경망의 학습과정에서 출력층의 오차를 입력층까지 충분히 전달하지 못하는 장기 의존성(long-term dependencies)의 문제 때문에 발생한다. LSTM(long short-term memory)은 순환 신경망의 장기 의존성의 문제를 극복하고자 하는 대표적인 모델이다(Hochreiter and Schmidhuber, 1997). 순환신경망을 이용한 감성분석 모델에 관한 연구로는 언어에 독립적으로 작동하는 감성분석 모델에 대한 연구가 시도된 바 있다(Tarasov, 2015). 또한 LSTM 구조를 활용한 감성분석에 대한 연구도 다양한 구조변형을 통해 시도되고 있는 상황이다(Tang et al., 2015).

### 3. 딥러닝 기반 감성분석 모델의 제안

본 연구에서는 감성대화챗봇 개발에서 대화자의 발화가 분노(anger), 죄책감(guilt), 두려움(fear), 혐오(disgust), 슬픔(sadness), 수치(shame) 및 기쁨(joy)의 일곱 가지 감성을 나타낸다고 보고 이들 발화를 텍스트로 바꾸어 발화내용을 일곱 가지 감성중의 하나로 분류하는 딥러닝모형을 개발하고자 한다.

<Figure 1>은 본 논문에서 제안하는 감성분석



<Figure 1> Deep Learning Process for Sentiment Analysis

을 위한 딥러닝 절차를 보여주고 있다. 제안하는 모형은 입력되는 임의의 문장단위의 텍스트 데이터에 대해서 word2vec을 이용하여 각 단어별로 특정 길이의 벡터로 임베딩 한 후 임베딩된 단어의 벡터를 딥러닝 기반의 심층신경망에 입력하여 결과 값으로 최종 감성을 분류한다.

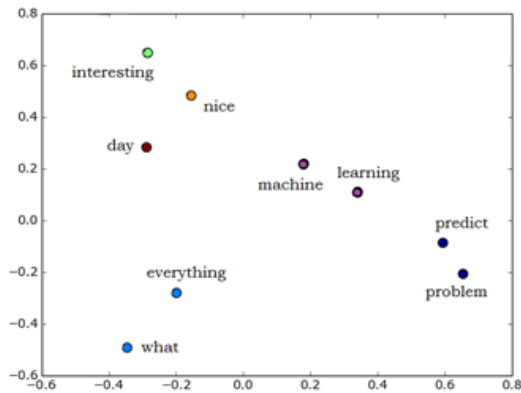
#### 3.1 워드 임베딩

텍스트 데이터를 이용하여 감성분석 모델을 구축하기 위해서는 먼저 주어진 데이터에 대한 전처리 과정이 필요하다. 전처리 과정에서는 문장의 감성을 분석하는데, 유의미한 정보를 제공하지 않는 마침표, 쉼표, 괄호, 구두점 등의 불용표현들이 워드 임베딩을 수행하기 전에 제거된다.

전처리 과정 이후에는 텍스트 데이터를 국소표현으로 변환한 뒤, 다시 word2vec 워드 임베딩 모델로 변환한다. 국소표현은 전체 단어집합의 원소의 수만큼 길이의 벡터를 통해, 해당하는 단어의 순번에만 1 값으로 그 이외의 경우에는 0 값으로 표현하는 방식을 의미한다.

다른 워드임베딩방법으로는 bag of words가 있는데 이는 문장을 단어의 집합으로 표현하여 그 집합 속에 있는 단어의 빈도를 벡터로 표현하는 방법이다. 본 연구에서는 실험의 비교를 위해 bag of words와 words2vec을 이용하였으며 실험데이터로 ISEAR(the International Survey of Emotion Antecedents and Reactions) 데이터 셋을 이용했다

(Scherer and Wallbott, 1994). ISEAR은 1990년대에 37개국 약 3,000명의 응답자로부터 얻은 7,350개 문장으로 구성된 텍스트 데이터와 그에 대한 감성을 정리한 설문조사 결과이다. 응답자는 문장이나 짧은 문장 또는 두 문장의 형태로 특정 감정을 느낀 상황이나 사건을 설명하도록 안내 받았으며, 데이터 셋의 감성 범주는 분노(Anger), 혐오감(Disgust), 두려움(Fear), 죄책감(Guilt), 기쁨(Joy), 슬픔(Sadness) 및 부끄러움(Shame) 등 모두 일곱 가지 감정으로 되어있다. 따라서 총 7,350개의 데이터를 40%는 훈련용, 나머지 40%는 검증용, 나머지 30%는 테스트데이터로 나누어 word2vec 워드임베딩으로 이를 200 차원 길이의 벡터공간의 벡터로 변환시켰다. 학습된 word2vec 모델을 이용하면 각 코퍼스는 이 벡터공간의 특정 지점으로 할당된다.

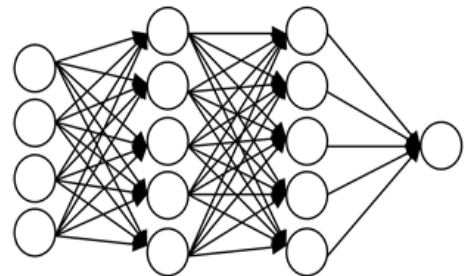


<Figure 2> Example of Word2vec Word Embedding

<Figure 2>는 word2vec을 통해 임베딩된 단어들을 2차원 공간으로 표현한 예시이다. 이러한 방법으로 임베딩된 단어들은 입력되는 코퍼스의 공통 내용에 따라서 유사도를 갖게 되어, 유사성이 있는 단어들끼리 가까운 거리를 갖도록 임베딩 시킬 수 있다. 예를 들어, <Figure 2>에서 'interesting'과 'nice'는 상대적으로 가까운 거리에 임베딩 되고, 'problem'은 상대적으로 먼 거리에 임베딩 되는 것을 확인할 수 있다.

### 3.2 딥러닝모델의 개발

다범주 감성분석을 수행하는 본 연구의 핵심은 발화내용의 감성을 분류하는 모델을 설계하고 학습시키는 것이다. 딥러닝모델은 다층신경망과 비교하여 은닉층의 수와 구조적인 측면에서 차이점이 있다. 본 연구에서는 감성분석을 효율적으로 수행할 수 있는 적절한 분류모델을 구성하기 위해, 다층신경망, 합성곱신경망(CNN)과 LSTM을 설계하고 분류성능을 비교하였다.

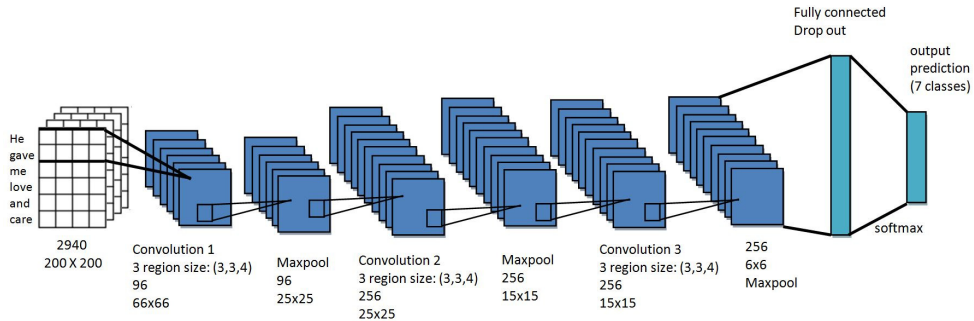


input layer hidden layer 1 hidden layer 2 output layer

<Figure 3> Example of Multi-Layer Neural Networks (MLNN)

첫 번째 분류모델은 다수의 은닉층으로 구성된 다층신경망(MLNN)을 이용한다. 다층신경망 모델에서는 2개의 은닉층을 갖는 모델을 설계하였으며 각 은닉층에서는 뉴런의 수를 50개부터 시작하여 50개씩 늘어나가 500개까지 실험하였다. 실험결과는 2개의 은닉층에 은닉노드가 500개인 모델 즉 200×500×500×7일 때 분류성능이 제일 양호한 결과를 얻었으며 실험결과는 <Table 1>과 같다.

두 번째 감성분석 모델은 합성곱신경망(CNN)이다. 합성곱신경망은 크게 합성곱층(convolution layer), 풀링층(pooling layer), 완전연결층(fully-connected layer)로 구성된 심층신경망을 의미한다. 합성곱신경망의 가장 큰 특징은 합성곱층과 풀링층을 통해 입력데이터의 특징을 추출하는 과정이다. 이러한 사전학습(pre-trained)과정을 통해 고차원의 데이터에서 특징들만 추출하여 저차원의 데이터를 얻을 수 있고, 이를 완전연결층에서 분류하는



<Figure 4> Proposed CNN Model

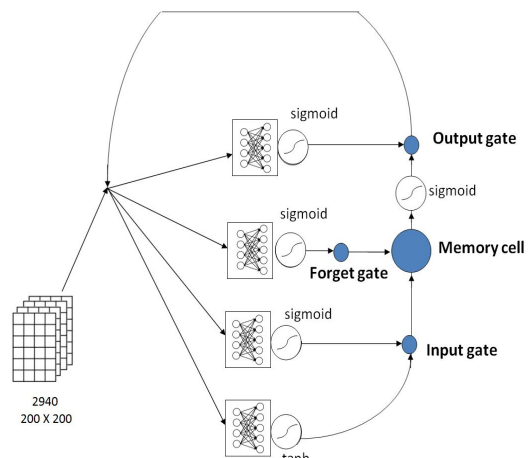
접근방식을 통해 분류모델로서 뛰어난 성능을 발휘하고 있다.

<Figure 4>는 임베딩 된 단어를 200×(훈련데이터의 수) (여기서 훈련데이터는 전체데이터 7,350개의 40%로 약 2,940개임) 크기의 2차원 행렬로 구축하고 최종적으로 일곱 개의 감성을 분류하는 합성곱신경망의 구조를 나타내고 있는데 세 개의 합성곱층과 하나의 완전연결층으로 구성되며, 각 합성곱층에서는 합성곱 연산과 맥스풀링 연산이 차례대로 수행되며 완전연결층의 이후에는 softmax 연산을 이용하여 최종적인 감성분석을 수행한다. 합성곱신경망의 입력 데이터는 텍스트 데이터에서 워드 임베딩을 통해 얻은 문장단위의 2차원 행렬을 입력값으로 한다. 워드임베딩을 통해 2차원의 입력 데이터가 준비되면 필터(filter)를 통해 입력데이터를 읽어 내려가며 입력 데이터와 필터와의 합성곱 연산을 수행하여 특징지도(feature map)를 생성한다.

필터 크기는 합성곱층에서 입력되는 값들에 대한 연산을 처리할 때, 한 번에 처리할 수 있는 이웃 정보의 크기라고 볼 수 있다. 사전에 정의된 입력 이미지의 크기가 결정되면, 합성곱층 필터의 최적의 크기는 실험적으로 결정하는데 본 연구에서 필터크기는 <Table 2>에서와 같이 다양한 크기로 하여 분류정확도가 높은 필터크기를 최종으로 결정하였다(Zeiler and Fergus, 2014). 합성곱신경망의 실험결과는 다음 절의 <Table 2>와 같다.

세 번째 감성분석 모델은 Long Short Term Memory(LSTM)이다. LSTM은 순환신경망의 일반

적인 문제점인 장기의존성의 문제를 해결할 수 있는 셀 구조로 구축된 순환신경망의 일종이다. LSTM은 순환신경망의 은닉층의 노드에 별도의 메모리 공간(cell)을 설정하고 그 메모리 공간으로 입력되거나 출력되는 길목에 특수한 게이트(gate)를 설치하여 장기의존성 문제에 접근하였다. LSTM에 적용된 게이트는 크게 세 가지로서 메모리 공간에 입력값을 저장할지 결정하는 입력 게이트(input gate)와 이전 상태값의 값을 얼마나 반영할지 결정하는 망각 게이트(forget gate), 메모리 공간의 출력을 결정하는 출력 게이트(output gate)이다. LSTM은 이와 같은 게이트를 통한 학습을 통해 입력데이터의 길이에 강인한 순환신경망 계열의 감성분석 모델로서 작동한다.



<Figure 5> Proposed LSTM Model

<Figure 5>는 본 연구에서 제안하는 LSTM 모델이며 세 개의 게이트와 하나의 셀로 구성된 LSTM의 은닉층 노드의 구조를 나타낸다. 이 구조 내에서는 입력 게이트, 망각 게이트, 셀 활성화, 출력 게이트 등의 순서로 연산되어 각 노드의 출력값이 산출된다. 최종 순서의 텍스트 데이터까지 LSTM에 입력되면, LSTM은 현재의 입력값과 더불어 그 이전까지의 모든 상태 값들의 영향력이 반영된 최종 출력값을 이용해 다범주 감성분석을 수행한다.

앞 절의 합성곱신경망의 실험에서 word2vec이 bag of words보다 양호한 결과를 얻었기 때문에 LSTM 모델의 실험에서는 word2vec의 결과만을 입력으로 하여 실험하였으며 실험결과는 <Table 3>과 같다.

## 4. 실험 결과

### 4.1 실험 방법

실험은 window10 운영체제에서 Tensorflow 1.0을 이용하여 진행되었으며, 딥러닝모델의 효율적인 학습을 위해 GTX1080 GPU를 이용하였다.

<Figure 6>은 데이터의 분리, 워드 임베딩, 분류모델의 학습 등의 실험진행의 전체적인 과정을 나타낸다.

실험의 첫 단계는 40%의 비율로 학습 데이터, 30% 비율로 타당성 데이터, 나머지 30%의 비율로

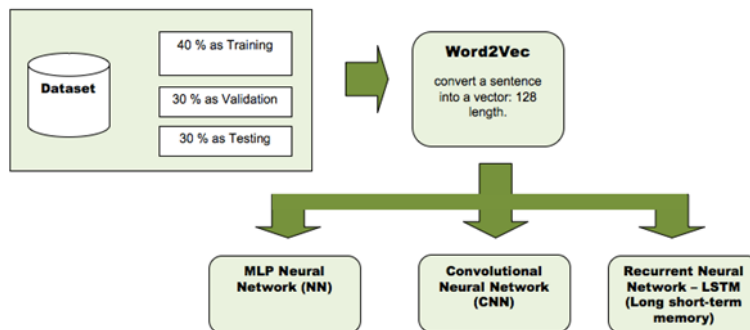
테스트 데이터로 나누고, 학습 데이터 셋은 딥러닝 기반의 심층신경망 분류모델의 학습에 사용하며, 학습의 과정에서 타당성 데이터를 통해 학습이 정상적으로 수행되는지 확인한다. 마지막으로 학습을 마친 모델에 대해서 테스트 데이터 셋을 이용하여 성능평가를 수행한다. 워드 임베딩 모델로는 bag of words와 word2vec을 사용했으며, 길이가 200인 벡터로 임베딩이 수행되도록 임베딩 모델을 구축하였다. 마지막으로 제안하는 감성분석 모형에서 사용한 심층신경망 분류모델은 일반적인 다층신경망, 합성곱신경망, LSTM 등 총 세 가지 종류의 신경망 모델을 이용하여 감성분석 모델의 성능 비교를 수행하였다.

특히, 합성곱신경망의 경우에는 입력 데이터의 특징을 추출하는 부분인 합성곱층의 필터의 크기를 조절하여 성능 변화를 파악했으며, LSTM의 경우에는 메모리 구조의 수를 변화시켜서 그에 따른 성능 변화를 측정하였다. 최종적으로는 각 모델별 분류 성능을 비교하였다.

이와 같은 성능비교를 위해 정밀도(precision) 및 재현율(recall)과 더불어 위의 두 값의 조화평균을 의미하는 F측도(F-measure)를 이용하여 성능평가를 진행하였다.

### 4.2 실험 결과

첫 번째 실험은 다층신경망(MLNN)을 이용하여 감성분석을 수행하였다. 이 실험에서는 다층신경망



<Figure 6> Three Models for Multi-Class Sentiment Analysis

의 은닉층의 수와 더불어 각 은닉층의 뉴런의 수까지 조절하여 성능비교를 수행하였다. <Table 1>은 다층신경망을 이용한 감성분석의 결과이며 테스트 데이터 셋을 이용한 분류정확도는 두 개의 은닉층에 노드수가 500개인 모델의 경우 분류정확도가 75.64%로 가장 높음을 보이고 있다.

<Table 1> Classification Accuracy of MLNN Model

number of hidden layers	number of nodes	accuracy (%)	F-1
1	50	56.00	0.58
	100	59.89	0.60
	200	61.72	0.60
	300	61.77	0.61
	400	62.00	0.62
	500	60.42	0.59
2	50	58.76	0.59
	100	62.59	0.63
	200	70.11	0.70
	300	72.77	0.71
	400	70.80	0.69
	500	75.64	0.75

두 번째 실험은 합성곱신경망을 이용한 실험에서는 합성곱신경망 모델의 각 합성곱층 마다의 필터 크기를 달리하였으며, bag of words와 word2vec 워드임베딩 기법을 개별적으로 적용하여 분류성능을 평가하였다. <Table 2>는 각 합성곱층의 필터의 크기에 따른 합성곱신경망 분류모델의 성능을 나타낸다. 필터의 조합에 따른 합성곱신경망 모델의 성능을 비교하면, 일반적으로 필터의 크기가 작을수록 보다 뛰어난 성능을 기록하는 경향을 확인할 수 있다. 또한 워드 임베딩 기법에 따른 성능비교를 통해 단어의 분산표현이 가능한 word2vec 기법을 사용하고 필터크기가 (4, 3, 4)일 때 분류정확도가 80.66%로 가장 우수함을 보이고 있다.

세 번째 실험은 순환신경망의 한 종류인 LSTM 모델을 이용하여, LSTM의 셀의 유닛 수를 30개에서 50개까지 10개씩 증가시킴으로써 그 성능을 비교하였다.

<Table 3>은 LSTM의 셀의 수 변화에 따른

<Table 2> Classification Accuracy of CNN Model

filter size	accuracy		F-1	
	bag of words	word2vec	bag of words	word2vec
(5, 5, 5)	74.51	75.00	0.73	0.73
(5, 5, 4)	62.41	64.44	0.63	0.65
(5, 5, 3)	57.11	57.11	0.57	0.59
(5, 4, 5)	77.8	78.00	0.77	0.78
(5, 4, 4)	63.18	64.00	0.64	0.67
(5, 4, 3)	64.00	71.88	0.64	0.70
(5, 3, 5)	67.55	64.88	0.66	0.65
(5, 3, 4)	63.89	64.22	0.64	0.66
(5, 3, 3)	71.50	72.66	0.71	0.72
(4, 4, 5)	64.88	77.77	0.65	0.76
(4, 4, 4)	71.70	71.77	0.72	0.72
(4, 4, 3)	55.32	58.55	0.55	0.59
(4, 3, 5)	56.66	56.66	0.55	0.55
(4, 3, 4)	76.71	80.66	0.76	0.79
(4, 3, 3)	65.80	67.88	0.65	0.67
(3, 5, 5)	57.00	57.66	0.57	0.57
(3, 5, 4)	71.00	71.00	0.71	0.71
(3, 5, 3)	71.00	71.11	0.71	0.71
(3, 4, 5)	73.40	73.44	0.72	0.72
(3, 4, 4)	65.11	65.11	0.66	0.67
(3, 4, 3)	74.88	75.00	0.73	0.74
(3, 3, 5)	75.00	75.11	0.76	0.81
(3, 3, 4)	76.40	80.44	0.75	0.82
(3, 3, 3)	77.01	66.66	0.76	0.71

<Table 3> Classification Accuracy of LSTM Model

number of units in LSTM cell	accuracy(%)	F-1
30	80.89	0.80
40	82.00	0.82
50	84.66	0.84
60	84.55	0.84

분류성능 변화를 나타낸 표이다. 실험결과, 상대적으로 많은 수의 단위 셀이 50개인 경우 분류정확도가 84.66%로 가장 우수한 성능을 보이며 F-1값도 양호하다. LSTM 셀의 유닛의 수는 일종의 연결가중치를 의미하는 것으로서 신경망의 정보를 저장하는 공간이 커질수록 상대적으로 성능이 향상됨을 확인할 수 있다.



앞서 실험한 다층신경망, 합성곱신경망, LSTM 모델 중에서 테스트 데이터에 대한 분류정확도가 가장 높은 결과만을 정리하면 아래의 <Table 4>와 같다.

<Table 4> Comparison of Accuracy for Three Models

	MLNN	CNN		LSTM
		bag of word	word2vec	
accuracy (%)	75.64	77.01	80.66	84.66
F-1	0.75	0.77	0.82	0.84

특히, 합성곱신경망의 경우, 워드 임베딩 기법을 달리하여 워드 임베딩에 따른 감성분석 성능의 차이를 비교하기도 하였다. 실험결과 다층신경망의 분류정확도는 75.6%로 가장 낮으며, 그 다음으로는 bag of word 워드 임베딩 기법을 사용한 합성곱신경망 모델이 77.01%, word2vec을 이용한 합성곱신경망은 80.66%, LSTM 모델이 84.66%로 LSTM 모델이 가장 우수한 분류정확도를 보이고 있다. <Table 5>는 LSTM 모델을 이용했을 때의 오분류표를 나타내고 있다.

마지막으로 LSTM을 이용한 감성분석 모델의 경우 84.79%로 가장 높은 분류 정확도를 보였으며, 딥러닝 기반의 심층신경망을 이용하여 총 일곱 개의 감성을 효율적으로 분류할 수 있음을 보였다. <Table 5>는 테스트데이터에 대하여 LSTM 모델을 적용했을 때의 오분류표를 나타내고 있다.

<Table 5>에서 알 수 있듯이 LSTM 모델이

모든 감성들을 대체적으로 잘 분류하고 있음을 나타낸다.

### 5. 결론과 향후연구

기존의 감성분석과 관련된 연구에서는 긍정과 부정 등의 이진 감성분석에 대한 연구들이 다수였다. 이에 본 연구에서는 사람의 대화에서 표현되는 감성, 즉, 기쁨, 슬픔, 두려움, 수치, 죄책감, 분노, 혐오 등 일곱 가지 감성을 분류할 수 있는 딥러닝모델을 개발하고 실증적으로 분류정확도가 높음을 보였다. 제안된 딥러닝모델은 고객의 상품평 분석, 은행, 보험회사에서 고객 상담 챗봇, 가정에서 도우미 챗봇등 감성대화가 필요한 챗봇에 활용 가능성이 매우 크다고 할 수 있다.

향후연구로는 발화자의 음성 텍스트 데이터뿐 아니라 음성, 얼굴 표정 같은 이종데이터를 결합하는 것이 보다 휴머노이드 감성대화 로봇개발의 나아갈 방향이 될 것이다.

### References

Bakliwal, A., J. Foster, J. Van der puil, R. O'Brien, L. Tounsi, and M. Hughes, "Sentiment Analysis of Political Tweets : Towards An Accurate Classifier", *Proceedings of the NAACL Workshop on Language Analysis in Social Media. Association for Computational Linguistics*, Atlanta, GA, 2013, 50-58.

<Table 5> Confusion Matrix Using LSTM

	Anger	Disgust	Fear	Guilt	Joy	Sadness	Shame
Anger	268	7	5	10	4	10	11
Disgust	13	269	9	5	3	9	7
Fear	13	3	264	8	6	12	9
Guilt	12	2	9	263	8	10	11
Joy	6	2	2	8	271	15	11
Sadness	13	2	3	3	15	270	9
Shame	11	8	2	15	5	15	259

- Bollen, J., H. Mao, and X. Zeng, "Twitter Mood Predicts The Stock Market", *Journal of Computational Science*, Vol.2, No.1, 2010, 1-8.
- Ding, X., B. Liu, and P.S. Yu, "A Holistic Lexicon-based Approach to Opinion Mining", *Proceedings of the 2008 International Conference on Web Search and Data Mining*, New York, USA, 2008, 231-240.
- Dos Santos, C.N. and M. Gatti, "Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts", *Proceedings of COLING, the 25<sup>th</sup> International Conference on Computational Linguistics*, Dublin, Ireland, 2014, 23-29.
- Feldman, R., "Techniques and Applications for Sentiment Analysis", *Communications of the ACM*, Vol.56, No.4, 2013, 82-89.
- Hochreiter, S. and J. Schmidhuber, "Long Short-term Memory", *Neural Computation*, Vol.9, No.8, 1997, 1735-1780.
- Kim, Y., "Convolutional Neural Networks for Sentence Classification", *arXiv Preprint arXiv : 1408.5882*, 2014.
- Krizhevsky, A., I. Sutskever, and G.E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks", *Proceedings of the 25<sup>th</sup> International Conference on Neural Information Processing Systems*, Vol.1, Nevada, USA, 2012, 1097-1105.
- Liu, B., *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers, USA, 2012.
- Liu, Y., J.W. Bi, and Z.P. Fan, "Multi-class Sentiment Classification : The Experimental Comparisons of Feature Selection and Machine Learning Algorithms", *Expert Systems with Applications*, Vol.80, No.1, 2017, 323-339.
- Medhat, W., A. Hassan, and H. Korashy, "Sentiment Analysis Algorithms and Applications : A Survey", *Ain Shams Engineering Journal*, Vol.5, No.4, 2014, 1093-1113.
- Mikolov, T., K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space", *arXiv Preprint arXiv : 1301.3781*, 2013.
- Mullen, T. and N. Collier, "Sentiment Analysis Using Support Vector Machines with Diverse Information Sources", *Processing of the Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, 2004, 412-418.
- Pang, B., L. Lee, and S. Vaithyanathan, "Thumbs up? : Sentiment Classification Using Machine Learning Techniques", *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Stroudsburg, USA, 2002, 79-86.
- Pang, B. and L. Lee, "Opinion Mining and Sentiment Analysis", *Foundations and Trends in Information Retrieval*, Vol.2, No.1-2, 2008, 1-135.
- Poria, S., E. Cambria, and A. Gelbukh, "Deep Convolutional Neural Network Textual Features and Multiple Kernel Learning for Utterance-level Multimodal Sentiment Analysis", *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Lisbon, Portugal, 2015, 2539-2544.
- Scherer, K.R. and H.G. Wallbott, "Evidence for Universality and Cultural Variation of Differential Emotion Response Patterning", *Journal of Personality and Social Psychology*, Vol.66, No.2, 1994, 310-338.

- Taboada, M., J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based Methods for Sentiment Analysis", *Computational Linguistics*, Vol.37, No.2, 2011, 267-307.
- Tai, K.S., R. Socher, and C.D. Manning, "Improved Semantic Representations from Tree-structured Long Short-term Memory Networks", *Proceedings of the 53<sup>rd</sup> Annual Meeting of the Association for Computational Linguistics and the 7<sup>th</sup> International Joint Conference on Natural Language Processing*, Beijing, China, 2015, 1556-1566.
- Tan, S., X. Cheng, Y. Wang, and H. Xu, "Adapting Naive Bayes to Domain Adaptation for Sentiment Analysis", *Advances in Information Retrieval, Lecture Notes in Computer Science*, Vol.5478, 2009, 337-349.
- Tang, D., B. Qin, X. Feng, and T. Liu, "Effective LSTMs for Target-Dependent Sentiment Classification", *arXiv Preprint arXiv : 1512.01100*, 2016.
- Tarasov, D., "Deep Recurrent Neural Networks for Multiple Language Aspect-based Sentiment Analysis of User Reviews", *Proceedings of the 21<sup>st</sup> International Conference on Computational Linguistics*, Kazan, Russia, 2015.
- Zeiler, M.D. and R. Fergus, "Visualizing and Understanding Convolutional Networks", *arXiv Preprint arXiv : 1311.2901*, 2013.

## ◆ About the Authors ◆



**M. Alex Syaekhoni (alex@dongguk.edu)**

M. Alex Syaekhoni got a Bachelor degree in Informatics Engineering in 2010, a Master degree in Industrial & Systems Engineering in 2013 and currently a doctoral student in Industrial & Systems Engineering, Dongguk University since 2013. His interests include machine learning, data mining, intelligent information systems, and customer relationship management.



**Sang Hyun Seo (shseo@dongguk.edu)**

Sang Hyun Seo has graduated from Dongguk University, Korean Language and Literature (BA) and a M.S. student in the Department of Computer Engineering at Dongguk University. His interests include machine learning, deep learning and reinforcement learning



**Young S. Kwon (yskwon@dongguk.edu)**

Young S. Kwon is a professor at Dongguk University. He graduated from Seoul National University and got his Ph.D degree from KAIST. His interests lies in machine learning and data mining applications.