

---

저자 (Authors)	류진걸, 신동민
출처 (Source)	<a href="#">대한산업공학회 추계학술대회 논문집</a> , 2014.11, 897-903(7 pages)
발행처 (Publisher)	<a href="#">대한산업공학회</a> Korean Institute Of Industrial Engineers
URL	<a href="http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE02511231">http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE02511231</a>
APA Style	류진걸, 신동민 (2014). SVM과 HCRF를 이용한 텍스트 문서 감정 분류 모델. 대한산업공학회 추계학술대회 논문집, 897-903
이용정보 (Accessed)	가천대학교 203.249.***.201 2019/09/29 19:06 (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.

# SVM과 HCRF를 이용한 텍스트 문서 감정 분류 모델

류진걸 신동민\*

한양대학교 산업경영공학과

---

본 연구는 국방과학연구소(계약번호 UD120033JD)의 연구비 지원 하에 인간 판단 모의기법 기술 과제의 일부로서 수행되었음.

\* 교신저자: 신동민 교수, 426-791 경기도 안산시 상록구 한양대로 55 한양대학교 산업경영공학과, Tel: 031-400-5268,

E-mail: dus122@gmail.com

## Abstract

본 논문은 SVM(Joachims)과 HCRF(Hidden Conditional Random Field)를 이용하여 텍스트 문서의 감정을 분류하는 시스템을 제안한다. 텍스트 문서 중에서도 SNS나 댓글과 같은 문서는 일반적인 문서와 달리 한두 문장 정도로 그 길이가 짧기 때문에 포함하고 있는 키워드가 많지 않다. 본 논문에서는 이와 같은 텍스트 문서의 특성을 반영하여, 문서에 포함된 키워드를 독립적으로 고려한 후, 인접 키워드들이 갖는 각 감정의 연관성을 고려하여 텍스트 문서의 감정을 분류하는 모델을 구현한다. 본 모델은 형태소로 분리되어 선별된 키워드를 자질로 하는 SVM을 학습시켜 얻은 자질의 계수를 HCRF의 입력 값으로 사용하며, HCRF를 이용하여 생성된 이 모델은 입력된 텍스트 문서를 긍정, 부정으로 분류한다. 영화 댓글 데이터를 수집하여 실험한 결과, SVM으로 구현된 분류기에 비해 높은 성능을 나타냈다. 향후 이 모델이 예 영화에 대한 댓글 뿐 아니라 상품 댓글이나 SNS상의 글 등의 감정분류에도 적용될 수 있을 것으로 기대된다.

## 1. 서론

오피니언 마이닝은 여러 형태의 텍스트 데이터에서 나타내는 의견 또는 감정을 추출하는 것을 의미한다. 오피니언 마이닝은 문서, 문장 등의 입력되는 텍스트의 형태에 따라서, 그리고 사실적 정보인지, 주관적인 감정이나 의견에 대한 텍스트인지에 따라 분류할 수 있고, 각 경우에 대한 다양한 연구가 이루어졌다. 특히, 최근에는 온라인 상에서 SNS, 영화 댓글 등의 단문의 데이터에 대한 감정을 분류하는 오피니언 마이닝 연구가 많이 이루어지고 있다(Go, Bhayani, & Huang, 2009; Kennedy & Inkpen, 2006; Pak & Paroubek, 2010;

Pang, Lee, & Vaithyanathan, 2002).

기계학습은 오피니언 마이닝에 사용되는 주요 방법 중 하나이다. 주로 사용되는 기계학습 기법에는 SVM, Naïve Bayes 등이 있으며, 많은 수의 자질과, 희소행렬 형태의 데이터에 대해 적합한 SVM이 일반적으로 높은 성능을 내는 것으로 알려져 있다(Joachims, 1998).

기계학습을 이용한 오피니언 마이닝의 연구분야는 크게 두 부분으로 나눌 수 있는데, 이는 효과적으로 텍스트 데이터를 분류하는 모델을 만드는 것과, 감정 분류의 정확도를 높일 수 있는 자질을 추출하는 것이다. 전자에 관한 연구로는 CRF(Conditional Random Fields)를 이용하여, 문장 간의 관계를 고려한 방법 등이 있으며, 후자에 관한 연구로는, 문장 내 단어의 패턴, 어휘와 품사의 패턴, 또는 품사로만 구성된 패턴 등을 이용한 자질 추출 방법 등이 있다(Cui, Mittal, & Datar, 2006; Fei, Liu, & Wu, 2004; Turney, 2002; Yang & Cardie, 2014).

감정분류의 정확성을 향상시키기 위해 많은 연구들이 이루어졌지만, 기계학습을 이용한 오피니언 마이닝은, 결국 특정 자질의 빈도수에 따라 그 자질의 성능이 달라지기 때문에 한계가 있다. 즉, 독립적인 단어나 자주 사용되는 패턴을 제외하면, 대부분의 패턴들은 감정을 분류하는데 큰 영향을 미치지 못하기 때문에, 특히 짧은 문장으로 이루어진 텍스트 데이터에서는 큰 힘을 발휘하지 못한다(Cambria, Schuller, Xia, & Havasi, 2013).

SNS, 영화 댓글 등은 한 두 문장으로 이루어진 단문이며, 문법을 따르지 않는 경우가 많고, 여러 감정 표현을 동시에 포함할 수 있다. 따라서 이러한 데이터는 문서나 문장에서 감정을 추출하는 것과는 다른 여러가지 문제를 갖고 있다. 다음 문장은 인터넷 포털에서 수집한 영화평이다.

*“내가 제일 좋아하는 영화는 봉준호 감독의 살인의 추억 우리나라에도 이런 감독이 있다는 사실에 충격을 받았고 감동했었다 그런데 설국열차정말*

살인의 추억 찍은 감독 맞는지 의심스러웠다  
개연성 없는 전개 병사들 말고는 앞칸 사람들  
아무도 커티스 일행을 공격않고 전진하게 지켜본다”

위 댓글을 보면, 앞부분에는 해당 감독에 대한 긍정적인 표현을, 그리고 이후로는 영화에 대한 비평을 하고 있다. 이 댓글을, 글에 포함된 단어의 패턴이나, 키워드를 가지고 분석할 경우, ‘제일’, ‘좋아하는’, ‘충격을’, ‘감동했었다’, ‘의심스러웠다’, ‘개연성 없는’ 등의 키워드를 통해, 극성의 정도를 파악하고 분류하게 되는데, ‘제일’, ‘좋아하는’ ‘감동했었다’ 의 키워드가 매우 강한 긍정이기 때문에 긍정으로 분류될 가능성이 크다. 이러한 문제를 해결하기 위해서는, 포함된 자질의 감정의 정도 외에도, 문장 전체 감정의 흐름을 봐야 한다. 이 문장은, 맨 앞쪽에 강한 긍정으로 시작해서, 뒷부분은 계속 부정을 나타내는데, 이러한 감정의 흐름을 분석에 반영할 수 있으면, 그 정확도를 향상시킬 수 있을 것이다.

본 연구에서는 기존 기계학습을 기반으로 한 오피니언 마이닝의 문제점을 보완하기 위해, 기계학습으로 학습된 자질의 가중치를 이용하여 인접키워드 간의 감정을 같이 고려하는 방법을 제시한다. 새롭게 제안하는 방법은, 기존에 사용되는 자질을 사용하여 SVM을 통해 학습시킨 후에 생성되는 자질의 가중치 벡터를 자질의 감정으로 정의하고, 그 값을 HCRF의 입력값으로 사용하여, 문장의 감정을 분류하는 것이다. 결과적으로, 제안하는 방법은, 문장에 포함된 자질이 갖고 있는 감정의 크기와 극성, 문장 전체의 감정의 흐름을 고려하여 문장의 감정을 분류한다.

제안한 방법을 평가하기 위하여 포털 사이트 Daum에서 추출한 15만개의 댓글 데이터를 바탕으로 하여, 기존에 사용되는 기법 중 성능이 높다고 알려진 SVM과 제안한 방법의 성능을 비교하였다. 각 방법에 대해, 같은 데이터와 같은

자질을 이용한 실험 결과를 제시한다.

본 논문의 구성은 다음과 같다. 우선 2장에서는, 본 연구에서 제안하는 새로운 방법을 적용하기 위한 절차와, 사용된 SVM-HCRF 모델에 대해서 설명한다. 3장에서는 실험결과를 제시하고 마지막으로 4장에서는 결론을 맺는다.

## 2. 제안 모델

본 연구의 SVM, HCRF를 이용한 텍스트 문서의 감정 분류 모델은 자질의 가중치 계산, HCRF 감정 분류 모델 부분으로 나누어볼 수 있다.

자질의 가중치 계산 단계에서는, SVM을 이용하여 각 자질에 대한 가중치를 계산한다. 계산 과정은, 먼저 선택된 자질을 바탕으로 텍스트 문서의 감정을 분류하는 분류기를 학습시킨 후, 학습된 SVM 분류기의  $w$ 벡터를 추출한다. 이때  $w$ 벡터는 각 자질에 대응하는 계수이다.

모델 생성단계에서는 HCRF를 이용하여 댓글의 감정을 분류하는 모델을 생성한다. 이 모델은 댓글에 포함된 자질의 가중치를 입력 받아, 각 가중치의 값, 그리고 인접한 가중치의 값 등을 고려하여 히든노드의 계수를 계산한다. 그리고 댓글에 포함된 가중치의 값과 계산된 히든노드의 패턴을 바탕으로 입력된 댓글을 긍정, 부정으로 분류한다.

### 2.1. 자질 가중치 계산

자질 가중치의 부호는 긍정, 부정의 방향을 나타내고 그 크기는 해당 감정에 치우친 정도를 나타낸다. 이 값은 댓글을 긍정, 부정으로 분류하는 SVM을 학습하여 생성된 모델의  $w$ 벡터를 추출하여 구할 수 있다.

SVM은 주어진 데이터를 분류하기 위해서 최대의 마진을 갖는 초평면을 탐색하는 기법이다(Gunn, 1998). SVM은  $(x_i, y_i)$  ( $i=1, \dots$ ,

n)형태의 데이터를 학습한다. 여기서  $x_i$  는 자질 수만큼을 요소로 갖는 벡터이며,  $y_i$  는 그 벡터에 대응하는  $\{1, -1\}$ 중 하나를 갖는 클래스 라벨이다. 예를 들어 본 논문에서  $x_i$  는 개별 영화 댓글에 포함된 자질을 벡터형태로 표현한 것이며,  $y_i$  에서 +1은 긍정, -1은 부정 영화평을 의미한다. SVM은 주어진 데이터로부터 다음 수식이 최소가 되는 매개변수 값을 찾는다.

$$\tau(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2 \quad (1)$$

$$\text{subject to } y_i((\mathbf{w} \cdot \mathbf{x}_i) + b) \geq 1, i = 1, \dots, n \quad (2)$$

여기서 구해진  $\mathbf{w}$  벡터는  $\mathbf{x}$  에 대한 가중치 벡터이다. 댓글  $\mathbf{x}$  에 포함된 각 자질에 대응하는  $\mathbf{w}$  벡터 요소의 값은 해당 자질이 긍정인지 부정인지에 따라 극성이 달라지고, 감정의 강도에 따라 그 값이 달라진다. 결국,  $\mathbf{w}$  벡터는 자질의 감정을 수치화한 값이라고 할 수 있다.

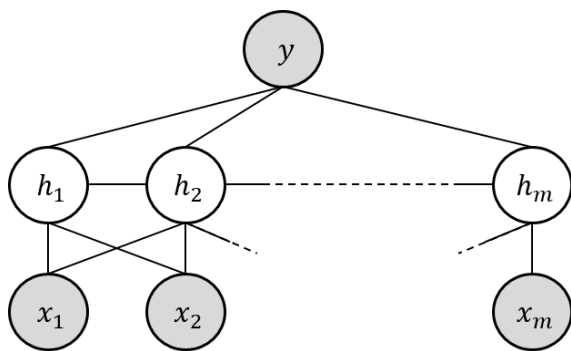


그림 1. 댓글의 감정 분류를 위한 HCRF 모델

## 2.2. HCRF 감정 분류 모델

HCRF는 기존의 CRF(Conditional Random Fields)에 hidden state를 도입하여, 입력 값의 잠재 구조를 구축할 수 있도록 한 방법이다(Quattoni, Collins, & Darrell, 2004). HCRF는 입력 값  $\mathbf{x}$ 를 받아,  $\mathbf{x}$ 에 대응하는  $\mathbf{h}$ 를 추정하고, 그를 바탕으로  $y$ 를 추정한다. HCRF의 조건부 확률 모델은 다음과 같이 정의된다.

$$p(y|\mathbf{x}, \theta) = \sum_{\mathbf{h}} P(y, \mathbf{h}|\mathbf{x}, \theta) = \frac{\sum_{\mathbf{h}} e^{\psi(y, \mathbf{h}, \mathbf{x}; \theta)}}{\sum_{y', \mathbf{h}} e^{\psi(y', \mathbf{h}, \mathbf{x}; \theta)}} \quad (3)$$

본 연구에서 제안하는 댓글의 감정 분류를 위한 HCRF 모델은 댓글  $\mathbf{x}$ 로부터 라벨  $y$ 를 추정한다.  $y$ 는 긍정, 또는 부정 둘 중에 하나의 값을 갖는다. 댓글  $\mathbf{x}$ 는 해당 댓글에 포함된 자질의 가중치  $\mathbf{x} = \{x_1, x_2, \dots, x_m\}$ 로 정의한다. 또한 모든  $\mathbf{x}$ 에 대해서,  $\mathbf{x} = \{x_1, x_2, \dots, x_m\}$ 에 대응하는 작성자의 감정  $\mathbf{h} = \{h_1, h_2, \dots, h_m\}$ 이 있다고 가정한다. 즉, 작성자의 감정을 표현하기 위해서는 하나의 단어가 아닌 여러 개로 이루어진 말뭉치가 필요하다고 가정하고, 그 감정을  $\mathbf{h}$ 로 표현한다.

댓글 감정 분류 모델을 위해 사용된 포텐셜 함수는 다음과 같다.

$$\psi(y, \mathbf{h}, \mathbf{x}; \theta) = \sum_j \phi(x_j) \cdot \theta(h_j) + \sum_j \theta(y, h_j) + \sum_{(j,k) \in E} \theta(y, h_j, h_k) \quad (4)$$

여기서  $\theta(k)$ 는  $k$ 번째 라벨에 대응하는 매개변수 벡터이다. 내적  $\phi(x_j) \cdot \theta(h_j)$ 는  $x_j$ 와  $h_j$ 의 호환성, 즉, 자질의 감정과 작성자의 감정이 얼마나 잘 맞는지를 나타낸다고 볼 수 있다.  $\theta(y, k)$ 는 작성자의 감정  $k$ 와 대응하는 댓글의 감정이 얼마나 잘 맞는지를 보여준다. 마지막으로  $\theta(y, k, l)$ 은 특정 댓글 전체의 감정  $y$ 와, 작성자의 감정  $k, l$  사이의 변화가 얼마나 잘 맞는지를 보여준다.

따라서 제안된 HCRF 감정 분류모델은 자질 가중치로부터 문장 내 키워드의 수치화된 감정을 입력 받아, 인접 키워드의 감정을 같이 고려하여 작성자의 감정을 추정한다. 결과적으로, 댓글에 포함된 키워드에 대해 추정한 감정과, 해당 키워드로부터 추정한 작성자의 감정, 그리고 그 변화에 따라 댓글의 최종 감정이 추정된다.

### 3. 실험

본 연구에서 제안하는 SVM, HCRF를 이용한 분류기의 성능과, SVM을 이용한 분류기의 성능을 비교하기 위해 영화 댓글을 긍정, 부정으로 분류하는 실험을 진행했다. 자질의 선택에 따라 그 성능이 많이 달라지는 문서분류기의 특성을 고려하여, 학습 집합과 테스트 집합에 포함된 긍정, 부정 문장의 비율을 달리하여 총 두 번의 실험을 진행했다.

#### 3.1. 데이터 구성 및 실험방법

수집된 데이터는 포털사이트 Daum의 네티즌평점이다. 영화명과 해당 영화에 대한 댓글과 평점, 추천수 등을 포함한다. 수집된 데이터에서 댓글의 내용과 평점이 일치하지 않는 데이터의 수를 줄이기 위하여 추천수가 3이상인 데이터만을 추출했다. 이 데이터 중 0, 1, 9, 10점의 평점을 갖는 리뷰와 평점만을 추출하였으며, 9, 10점 평점을 갖는 댓글을 긍정, 0, 1점 평점을 갖는 댓글을 부정으로 분류했다. 전처리를 거쳐 최종적으로 분석에 사용된 전체 댓글의 수는 총 22,056개이며, 이 중 16,761개가 긍정, 5,295개가 부정 댓글이다.

영화의 댓글을 긍정, 부정으로 분류하는 실험을 진행하였으며, SVM을 이용한 모델의 성능과, SVM, HCRF를 이용한 모델의 성능을 비교했다. 분석에 사용된 자질은 형태소 단위로 선정하였으며, 선정된 형태소는 외국어, 체언, 용언, 수식언, 독립언이다. 데이터는 긍정:부정 비율을 5:5로 맞춘 데이터와 실제 데이터의 긍정:부정 비율인 7:3으로 맞춘 데이터를 이용하여 각 데이터에 대한 성능을 비교한다. 성능 평가 방법으로는 정확률(Precision)과 재현율(Recall), 그리고 macro F-measure를 이용한다.

### 3.2. 실험 결과

실험 결과, SVM-HCRF를 사용할 경우 표 1. 긍정,

부정 비율이 1:1인 데이터에 대한 실험 결과

	SVM	SVM-HCRF
긍정 문장 정확률	0.89	0.89
부정 문장 정확률	0.80	0.83
긍정 문장 재현률	0.77	0.82
부정 문장 재현률	0.90	0.90
Macro F-Measure	0.84	0.86

표 2. 긍정, 부정 비율이 3:1인 데이터에 대한 실험 결과

	SVM	SVM-HCRF
긍정 문장 정확률	0.88	0.90
부정 문장 정확률	0.65	0.83
긍정 문장 재현률	0.89	0.96
부정 문장 재현률	0.61	0.68
Macro F-Measure	0.76	0.84

SVM만을 사용할 경우보다 그 정확성이 향상되는 것을 확인했다. 실제 데이터를 살펴봤을 때, ‘영화라는 장르에 대해 제대로 된 이해가 필요하다 영화로 만들었다면 영화다워야 한다 시종일관 답답하고 지루하다’ 라는 문장은 SVM에서만 오분류가 되었는데, 그 이유는, 이 문장에 포함된 ‘영화’, ‘장르’, ‘대해’, ‘제대로’, ‘되’, ‘이해’, ‘필요’, ‘영화’, ‘만들’, ‘영화’, ‘시종일관’, ‘답답’, ‘지루’의 자질 중, ‘이해’, ‘필요’가 다른 자질에 비해 월등히 높은 가중치를 갖고 있어서, 가중치의 합이 긍정을 나타냈기 때문이다. HCRF를 적용한 모델은, 단순히 가중치의 값만 이용해서 분류하는 것이 아니라, 해당 가중치를 이용하여 추정된 사용자의 감정, 그리고 그 감정의 변화 또한 고려하기 때문에 결과적으로 이 문장을 부정으로 분류할 수 있었던 것으로 판단된다. 또한 표 1과 표 2를 비교해 보면, 긍정, 부정 비율을 임의로 1:1로 조정하였을 경우, SVM과 SVM-HCRF의 성능차이가

크지 않았지만, 그 비율이 3:1인 데이터로 학습하였을 경우, 그 차이가 두드러지게 나타났다.

#### 4. 결론 및 추후연구

본 논문에서는 텍스트 문서 내 포함된 키워드의 감정을 독립적으로 고려한 후, 인접 키워드의 감정을 바탕으로 문서 전체 감정의 흐름을 추정한 뒤, 문서에 포함된 키워드의 감정의 정도와, 전체 감정의 흐름을 고려하여 텍스트 문서의 감정을 분류하는 모델을 제안했다. 영화 댓글 데이터를 이용한 실험 결과, 본 논문에서 제안하는 SVM-HCRF를 이용한 모델은 기존의 SVM보다 높은 성능을 보였다. 또한 이 모델은, 기존의 연구들과 독립적으로 사용되는 것이 아닌, SVM을 이용하여 개발된 모델의 성능을 보완하는 모델이라는 점에서 그 의미가 크다고 할 수 있다.

한편, 본 논문에서는 외국어, 체언, 용언 등의 몇몇 형태소를 자질로 하여 실험을 진행하였다. 최근 많은 연구들이 분류기의 정확성을 높이기 위한 여러 자질들을 연구하고 있는데, 이렇게 보다 높은 성능을 보이는 자질을 분석에 사용하여, 해당 자질에 매겨진 가중치를 SVM-HCRF모델에 적용시킨다면 보다 높은 성능의 감정 분류기를 만들 수 있을 것이라 생각된다.

## 참고문헌

- Cambria, Erik, Schuller, Bjorn, Xia, Yunqing, & Havasi, Catherine. (2013), New Avenues in Opinion Mining and Sentiment Analysis, *IEEE Intelligent Systems*, **28**(2), 15-21.
- Cui, Hang, Mittal, Vibhu, & Datar, Mayur. (2006), Comparative experiments on sentiment classification for online product reviews, *In Proceedings of the 21st National Conference on Artificial Intelligence*, 1265-1270
- Fei, Zhongchao, Liu, Jian, & Wu, Gengfeng. (2004), Sentiment classification using phrase patterns, *In Proceedings of the Fourth International Conference on Computer and Information Technology*, 1147-1152.
- Go, Alec, Bhayani, Richa, & Huang, Lei. (2009), Twitter sentiment classification using distant supervision, *CS224N Project Report, Stanford*, 1-12.
- Gunn, Steve R. (1998), Support Vector Machines for Classification and Regression *ISIS technical report*, **14**.
- Joachims, T. (1998), Text Categorization with Support Vector Machines: Learning with Many Relevant Features, *In Proceedings of the European Conference on Machine Learning*, 137-142.
- Kennedy, Alistair, & Inkpen, Diana. (2006), Sentiment classification of movie reviews using contextual valence shifters. *Computational Intelligence*, **22**(2), 110-125.
- Pak, Alexander, & Paroubek, Patrick. (2010), Twitter as a Corpus for Sentiment Analysis and Opinion Mining, *In Proceeding of the Seventh International Conference on Language Resources and Evaluation*, Valletta, Malta.
- Pang, Bo, Lee, Lillian, & Vaithyanathan, Shivakumar. (2002), Thumbs up?: sentiment classification using machine learning techniques, *In Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, **10**, 79-86.
- Quattoni, Ariadna, Collins, Michael, & Darrell, Trevor. (2004), Conditional random fields for object recognition, *In Advances in Neural Information Processing Systems*, **17**, 1097-1104
- Turney, Peter D. (2002), Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews, *In Proceedings of the 40th annual meeting on association for computational linguistics*, 417-424
- Yang, Bishan, & Cardie, Claire. (2014), Context-aware Learning for Sentence-level Sentiment Analysis with Posterior Regularization, *In Proceedings of ACL*.