

RESEARCH ARTICLE

Sentiment Analysis as a Predictor for Scores of Persuasive Essays

Dongkwang Shin
Gwangju National University of Education

설득적 에세이 점수의 예측도구로서의 감성분석 기법

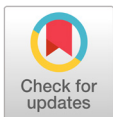
신동광
광주교육대학교

Corresponding Author: sdhera@gmail.com

ABSTRACT

This study aimed to investigate to what extent sentiment analysis can be applied to tone analysis of persuasive essays as a predictor for scores of the essays. To this end, the sentiment analysis of the text mining program ‘Orange3’ was utilized. This study first looked at the effects of Vader’s sentiment indices (positive, negative, neutral, and compound values) on the total score of positive and negative essays. As a result, in the positive essay, all the sentiment analysis results had a significant effect on the total score. In particular, it was confirmed that the compound value had the greatest influence as a predictor of the total score of the positive essay as well as the scores in the four sub-scoring domains—task performance, content, organization and language use. On the other hand, it turned out that there were no significant relationship between sentiment indices and scores of negative essays. These results suggest that even if learners write a persuasive essay with either positive or negative tone, they could expect a higher score when comparing the two opposite positions rather than the one-sided one. This was also confirmed to some extent through word cloud analysis. Of course, there were some limitations in this study as well but it showed the applicability of sentiment analysis to predict scores of persuasive essays.

Key words: Sentiment analysis, Orange3, persuasive essay, argument mining, word cloud



OPEN ACCESS

Brain, Digital, & Learning
2021, Vol. 11, No. 2, 215-226.
<https://doi.org/10.31216/BDL.20210014>

Received: May 04, 2021

Revised: May 26, 2021

Accepted: May 26, 2021

© 2021. Institute of Brain based Education,
Korea National University of Education



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

최근의 빅데이터 분석 기법의 하나로 감성분석이 다양한 분야에서 적용되고 있다(Shin, 2020). 이러한 감성분석 기법은 적용하는 분야나 목적에 따라 오피니언 마이닝(opinion mining)으로 불리기도 하고(Liu et al., 2005) 분석 데이터의 논증 방식의 타당성 분석에 적용될 때는 논증구조 마이닝(argumentation/argument mining) 등으로 불리기도 한다(Stab & Gurevych, 2017).

오피니언 마이닝의 경우, 페이스북이나 트위터와 같은 SNS나 게시판 등에 올라온 상품평 또는 특정 사안에 대한 대중의 반응을 긍정, 중립, 부정 등으로 분석하여 전반적인 여론의 경향성을 분석하는 과정을 의미한다. 반면 논증구조 마이닝은 설득적 에세이(persuasive essay)와 같이 논증구조가 중요한 영역에서 이를 기계적으로 분석하는 데 주로 활용되고 있다. 뿐만 아니라 이러한 감성분석 기법은 교육 분야에도 접목되고 있는 추세이다. 이는 대면교육에서 비대면 온라인교육으로 그 추세가 바뀌고 있는 현 시점에서 매우 중요한 시사점을 제공한다. 감성분석 기법이 탑재된 온라인 교육 시스템에서는 감성분석 기능이 학습자의 감정 상태를 분석하여 학습의 장애 요소를 파악하거나 감성적 피드백을 제공하여 학습 주의력을 높이는 데 활용되기도 한다(Arguedas et al., 2018). 이와 같이, 감성분석은 다양한 방식의 에듀테크로 진화하고 있다.

이중 본 연구에서 주목하고자 하는 분야는 논증구조 마이닝이다. TOEFL이나 IELTS와 같이 에세이 평가가 포함된 시험에서는 대부분 쓰기 평가 시 찬반 또는 장단점의 상반된 입장 중 한 가지 입장을 취하여 글을 작성하는 소위 설득적 에세이 문항을 포함하고 있다. 이와 같은 설득적 에세이에서는 글쓴이가 선택한 입장에 따라 긍정 또는 부정의 논조(tone/voice)를 가지고 글이 전개될 수밖에 없다. 이러한 관점에서 본 연구는 감성분석을 적용하여 설득적 에세이의 논조를 분석하면 그 감성분석 결과가 쓰기 채점의 특정 영역 또는 최종 점수를 예측하는 변인으로 활용 가능하리라는 가정 하에 시작되었다. 이를 검증하기 위해 본 연구는 다음과 같은 연구 질문을 설정하였다.

첫째, 긍정적 입장을 취한 에세이의 감성분석 결과에서는 어떠한 감성 유형이 각 채점영역(과제수행, 내용, 구성, 언어사용) 및 총점의 예측 변인으로 가장 타당한가?

둘째, 부정의 입장을 취한 에세이의 감성분석 결과에서는 어떤 감성 유형이 각 채점영역 및 총점의 예측 변인으로 가장 타당한가?

셋째, 두 입장의 에세이에 나타난 어휘 분포에는 어떠한 차이가 있는가?

Theoretical Background

Sentiment Analysis in the Field of Education

지금까지 여러 감성분석 알고리즘이 개발되었지만 크게는 감성사전을 활용하여 감정 상태를 분석하는 어휘기반 기법(lexicon based approach)과 빅데이터에 포함된 감성정보를 학습시켜 분석하는 기계학습 기법(machine-learning approach)으로 구분된다. Zhou & Ye (2020)에 따르면 최근의 감성분석 연구는 이 두 기법을 함께 적용하는 방식이 일반적이라고 보고되고 있다.

Ortigosa et al. (2014)의 연구에서는 감성분석 기법을 온라인 학습 시스템에 접목하여 학습자들의 반응을 실시간으로 분석하고 그 결과에 따라 적절한 피드백을 자동으로 제공하는 방식이 온라인 학습 시스템에 대한

학습자의 만족도 및 학력 증진에 도움이 된다는 것을 보여주었다. 이와 유사하게, Lin et al. (2014)의 연구에서는 학습 중 감성분석에 기반한 가상의 캐릭터를 활용하여 학습자들에게 피드백을 제공하였다. 학습자들은 이를 통해 학습 시스템과 상호작용하고 있다고 느낄 수 있었고 학습 결과에도 긍정적인 영향을 주는 것으로 나타났다.

Alm & Nkomo (2020)의 연구는 언어교육에 감성분석 기법을 사용한 첫 사례로 파악된다. Alm & Nkomo은 챗봇 기능을 포함하고 있는 온라인 언어 학습 애플리케이션(Duolingo, Eggfun, Memrise, Mondly)의 사용 후기를 분석하여 애플리케이션의 사용 만족도를 분석하였다. 이 분석에는 Microsoft Azure Cognitive Services의 감성분석기를 활용하였다. 분석 결과를 살펴보면, 사용자들은 어느 정도 챗봇 대화에 참여하고자 하는 의지를 보였고 참여 동기의 상당 부분은 챗봇에 대한 호기심에서 기인한 것으로 나타났다. 하지만 챗봇과의 대화가 학습내용과 일치하지 않거나 사용 중 발생한 오류에 대해서는 부정적인 반응을 보였다.

국내 연구로는 Shin (2020)의 연구에서 수능 영어영역의 문항 유형 중 하나인 ‘심경·분위기 파악’의 정답 예측에 감성분석 기법을 적용하였다. 어휘기반 기법과 기계학습 기법을 모두 적용한 이 연구에서는 기계학습에 기반한 IBM의 감성분석기 ‘Tone Analyzer’가 텍스트 마이닝(text mining) 프로그램 중 하나인 ‘KH-Coder’에 감성사전을 탑재하여 분석한 방식과 비교하여 더 높은 정답 예측률을 보였다. 하지만 두 개의 방식 모두 감정을 나타내는 어휘가 지문에 명시적으로 나타날 경우만 예측률이 높았고 맥락적 추론을 통해 화자의 심경이나 분위기를 파악해야 하는 문항에서는 매우 낮은 예측률을 보였다.

Persuasive Writing and Argument Mining

지난 수십 년간 쓰기 교육에서는 전통적인 모델에 기반하여 네 가지 스타일 즉, 묘사(description), 서술(narration), 설명(exposition), 논증(argument)과 관련한 쓰기 지도를 하거나(Braddock et al., 1963), 보다 기초적인 수준에서 단어, 문장, 또는 문단 단위에서 요구되는 특정 표현을 학습자가 작성 할 수 있도록 지도하는 데 초점을 두었다. 그 중에서도 논증적 또는 설득적 에세이는 논란이 있는 주제에 대하여 글쓴이가 한 가지 입장을 선택하여 독자를 설득하는 데 초점을 두는 쓰기 활동이다(van Eemeren et al., 1996).

Gentry (2010)에 따르면 학습자가 쓰기에 보다 능숙해 질수록 자신의 글에 대한 확신을 가지고 그에 따라 자신만의 논조를 명확히 반영하는 경향이 있다고 주장하였다. 이와 같은 맥락에서 Fowler (2019)는 설득적 에세이 작성을 위해 객관적인 정보를 통합하여 학습자 자신의 주장을 강화하도록 반영하고 독자들에게 인상적인 결론을 제시할 수 있도록 적절한 통사적 구조를 구사하는 것이 중요하다고 주장하였다.

Wilson (2008)은 설득적 에세이 작성에서 의견 제시에 사용되는 표현 유형을 두 가지로 분류하였다. 첫째는 감성 표현(sentiment expressions)으로 글쓴이의 판단, 평가 및 감정을 나타내며 특정한 입장을 지지하거나 이에 대한 반대 입장이 가진 문제점을 지적하는 데 사용된다. Table 1에 제시된 첫 번째 문장에서 “learned the most”는 글쓴이의 긍정적 평가를 명시적으로 보여주는 감성 표현이라고 할 수 있다. 두 번째의 의견 제시 유형은 논증 표현(arguing expressions)으로 그 다음 예문에 나타난 “For these reasons, I claim with confidence”에서는 글쓴이가 취하고 있는 입장을 근거와 함께 명확히 제시하고 있음을 확인할 수 있다. 이러한 표현은 독자를 설득하는 데 유용하며 에세이가 보다 설득력 있게 보이게 함으로써 그러한 표현이 없는 에세이보다 높은 점수를 받을 확률이 크다(Farra et al., 2015).

Table 1. Some expressions related to the two types of opinions (Farra et al., p. 66)

Types of Opinions	Sample Expressions
Sentiment	At school, I always <i>learned the most</i> from [teachers who encouraged me].
Arguing	<i>For these reasons, I claim with confidence</i> that [excellent knowledge of the subject being taught is secondary to the teacher's ability to relate well with their students].

최근에는 논증구조 마이닝 기법을 활용하여 설득적 에세이에서 글쓴이의 입장을 기계적으로 분석하려는 연구가 시도되고 있다. 다음에 제시된 논증구조의 분석 과정은 Stab & Gurevych (2017)가 제안한 분석 모델이다.

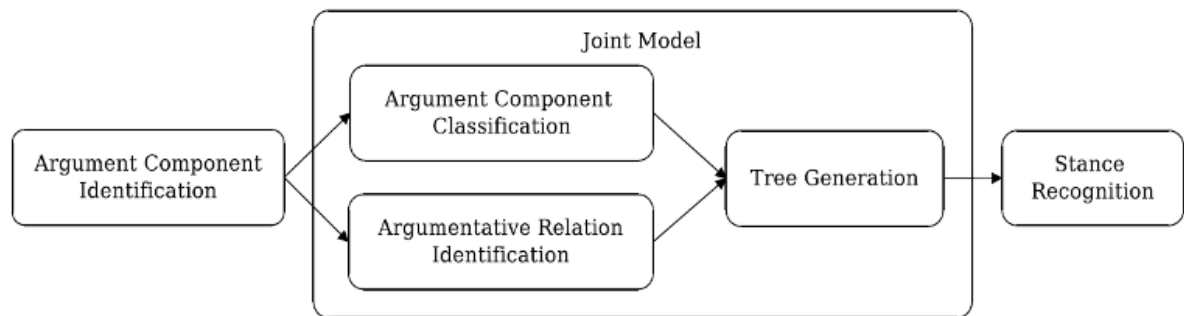
**Fig. 1.** Architecture of the argumentation structure parser (Stab & Gurevych, 2017, p. 635)

Fig. 1에 제시된 바와 같이, Stab & Gurevych (2017)가 제시한 모델에서는 논증구조의 분석을 위해 5개의 연속적인 하위 과업이 요구된다. 첫 번째의 구분(identification) 모델은 분석할 텍스트에서 논증 표현과 그 외의 표현을 구분한다. 그 다음에 이어지는 3개의 연계된 모델(Joint Model)은 각 단계의 세부 언어 자질들을 분석하여 논증구조를 인식한다. 논증 요인 분류기(Argument Component Classification)와 논증 관계 구분기(Argument Relation Identification)에서 산출된 결과는 각 문단별로 주장의 근거 또는 반박이라는 두 입장을 기준으로 다시 분류되며 전체적인 글쓴이의 입장을 확정하게 된다. 이러한 분석 절차 가운데 포함된 감성분석에서는 감성사전을 기반으로 한 단어별 감성분석과 문장별 감성분석의 과정이 포함되어 있다. 하지만 이러한 시도에도 불구하고 Stab & Gurevych는 에세이의 전반적인 점수를 예측하기 위해 의견이나 감성 정보를 활용하는 연구는 여전히 미연구 분야로 남아있다고 지적하였다.

Methods

Data Collection

본 연구의 참여자는 2020년 서울 소재 A 고등학교 4개 반 120명의 학생들이었다. 반별로 참여자를 등분하여 절반은 긍정, 나머지 절반은 부정의 에세이를 작성하게 하였다. 결과적으로, 긍정 입장의 에세이 60편과 부정 입장의 에세이 60편이 본 연구의 분석 대상이었다. 다음의 Table 2는 실험에서 활용한 설득형 에세이 문항이다.

Table 2. Writing prompt

On the advantages and disadvantages of having plastic surgery, choose one position, write an introduction, discuss with three reasons, and write a conclusion. (100-120 words)

Analysis Tools and Procedure

Persuasive Essay Scoring

120개의 답안에 대한 채점은 국가단위 시험에서 쓰기 채점에 경험이 풍부한 2명의 채점자가 참여하여 수행하였다. 채점영역은 Shin et al. (2016)의 서·논술형 수행평가 매뉴얼에 포함된 채점 기준표를 참고하였다. 채점영역은 과제수행, 내용, 구성, 언어사용이며 각 채점영역별 채점요인은 Table 3과 같다. 채점에는 채점불가인 0점을 제외한 1-5점 척도를 적용하였고 최종 점수는 4개 영역의 최대 점수는 5점을 합산한 20점을 만점으로 하였다. 또한 두 채점자간 신뢰도는 0.91로 높은 채점자 간 상관을 보였다.

Table 3. Scoring domains

Dimensions	Required Conditions
Task Performance (1-5)	-Completed an argumentative essay with 100-120 words -Explained the writer's position with three reasons -Explained the writer's position including the three parts of Introduction, Body, and Conclusion
Content (1-5)	-Clear main idea -Provided adequate supporting details in a logical manner -In-depth discussion of the main idea
Organization (1-5)	-Clear topic sentences -Clear supporting details -Clear concluding sentences -Logical progression of the main idea
Language Use (1-5)	-Used only English and appropriate vocabulary -Variety in sentence structures -Few grammar errors -Few errors in spelling -Appropriate use of mechanics

Sentiment Analysis

본 연구에서 설득적 에세이의 논조분석에 적용된 감성분석에는 Orange3 (<https://orangedatamining.com/download/#windows>)라는 프로그램이 활용되었다. Demsar et al. (2013)이 개발한 Orange3는 데이터 시각화, 기계 학습 및 데이터 마이닝 분석 서비스를 제공하는 무료 프로그램이다. Orange3의 특징 중 하나는 모든 기능이 포함된 패키지를 다운로드 받는 것이 아니라 기본 프로그램에 추가 기능이 필요할 때마다 다운로드 받는 방식을 채택하고 있다. 기본 옵션에는 감성분석과 같은 텍스트 마이닝(Text Mining) 기능이 포함되어 있지 않아 이 기능을 추가로 다운로드 받아야 한다. 프로그램 상단의 'Options'에서 'Add-ons'를 선택하고 'Orange3-Text'를 다운로드 받으면 좌측 기능 아이콘에 텍스트 마이닝 분석 기능이 추가된다.

일반적으로 R 프로그램과 같은 텍스트 마이닝 지원 소프트웨어는 사용자가 코딩 작업을 해야 하는 번거로움이 있으나 Orange3의 경우는 코딩 과정이 전혀 요구되지 않아 코딩에 대한 지식이 없는 사용자도 쉽게 활용할 수 있다는 장점이 있다. 엑셀 파일에 Fig. 2와 같이 ‘ESSAY’와 ‘NO.’라는 칼럼명을 지정한 후 분석 데이터를 입력하면 분석 데이터의 준비는 끝난다. 분석 절차로, 좌측 분석 메뉴에서 ‘Text Mining’을 선택한 후 ‘Corpus’와 ‘Sentiment Analysis’ 아이콘을 마우스로 끌어다 화면 중앙에 배치한다. 그리고 메뉴에서 ‘Data’를 선택하고 ‘Data Table’이라는 아이콘을 추가로 배치한다. 마우스로 이 3개의 아이콘을 Fig. 2의 하단 그림처럼 줄로 이어준다. ‘Corpus’ 아이콘을 더블 클릭한 후 분석 데이터를 업로드하고 ‘Sentiment Analysis’에서는 감성분석의 알고리즘 중 하나인 ‘Vader’를 선택해 준다. ‘Vader’는 Hutto & Gilbert (2014)가 구축한 ‘Vader’라는 감성사전을 이용하는 기법으로 세밀한 감성분석이 가능하다(Jhang et al., 2019). 마지막으로 ‘Data Table’을 더블 클릭하면 분석 결과를 확인할 수 있다. ‘Vader’ 감성분석은 긍정(positive), 부정(negative), 중립(neutral) 값과 더불어 통합(compound) 값을 제공한다. 긍정, 부정, 중립의 감정은 +1에 가까울수록 해당 감정의 강도가 세다는 것을 의미하고 통합 값의 경우 +1에 가까우면 긍정, 0에 가까우면 중립, -1에 가까우면 부정의 감정을 의미한다.

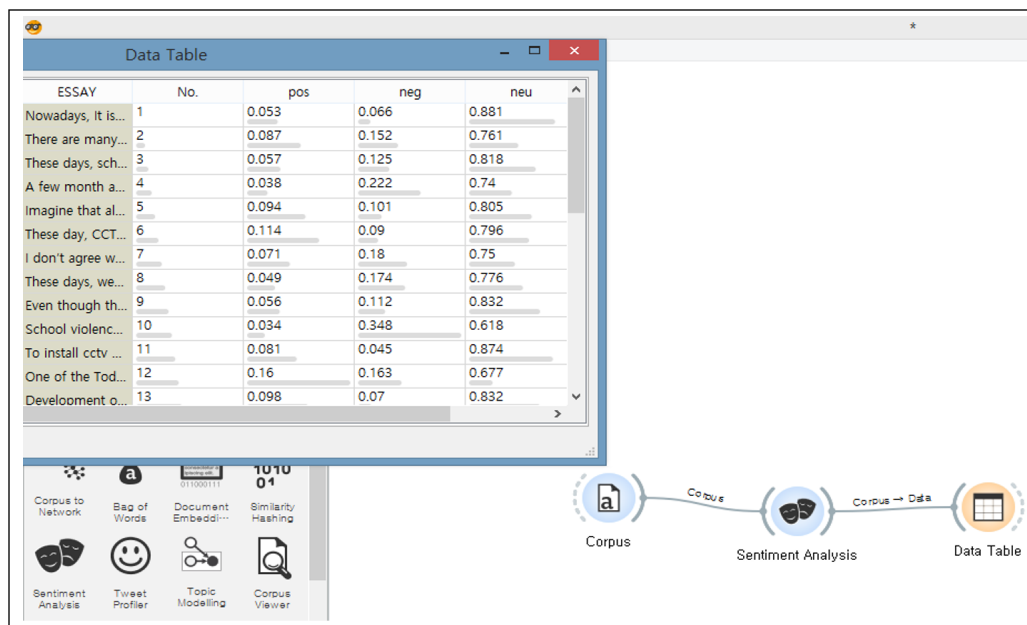


Fig. 2. Sentiment analysis on Orange3

Word Cloud Analysis

본 연구에서는 감성분석에 더하여 긍정과 부정의 입장을 취한 에세이를 구성하는 고빈도 단어를 추출하였다. 이를 위해 먼저 Orange3의 워드클라우드 분석 기능을 활용하였다. 분석 절차는 Fig. 3과 같이 ‘Word Cloud’와 ‘Preprocess Text(전처리)’ 절차를 추가 배치하고 각 아이콘을 선으로 연결하면 된다. 전처리 절차에서는 아래 Fig. 3 하단에 제시된 바와 같이 기능어(예, the, a, of 등) 목록을 텍스트 파일로 ‘Stopwords’에 지정하면 된다. 목록에 포함된 기능어는 분석에서 제외되기 때문에 내용어 중심으로 고빈도 단어의 추출이 가능하다. 이 분석의 ‘Stopwords’에 추가된 기능어는 총 275개 단어이며 ‘can’과 같은 조동사는 설득적 에세이에서 중요한 의미를 전달할 수 있어 ‘Stopwords’에서 제외되었고 ‘not’과 같은 부정어로 같은 이유로 제외되었다.

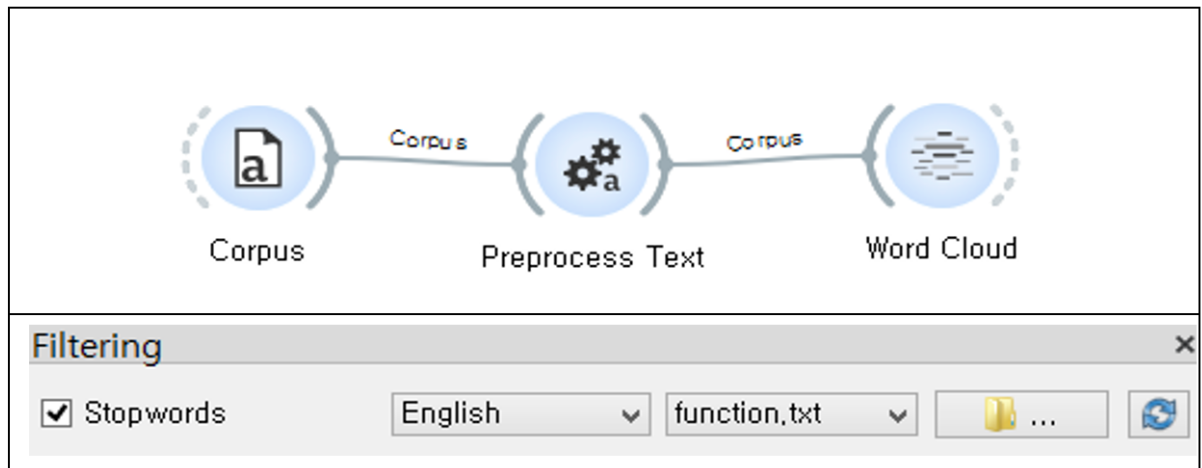


Fig. 3. Word cloud analysis on Orange3

Results and Discussion

Sentiment Analysis as a Predictor for Scores of the Positive Essay

본 연구에서는 앞서 연구 질문에서 제시한 바와 같이 긍정과 부정의 설득적 에세이의 점수 예측을 위한 변인으로서 감성분석 결과를 살펴보았다. 먼저 긍정적 에세이에 대한 감성분석의 기술통계치는 다음의 Table 4와 같으며 이 4개 변인들(긍정, 부정, 중립, 통합 값)이 에세이 평가의 총점에 미치는 영향을 측정하기 위한 다중회귀분석(Multiple Regression Analysis)의 결과는 Table 5와 같다.

Table 4. Descriptive statistics of positive essays

Variables	M	SD	n
Total Score	10.92	3.61	60
Positive	0.23	0.09	60
Negative	0.05	0.04	60
Neutral	0.72	0.10	60
Compound	0.78	0.31	60

Table 5. Multiple regression analysis on sentimental factors as a predictor for the positive essay score

Model	Predictor Variables	Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	β	t	Sig.
Positive Essay	(Constant)	6.69	1.72		3.89	0.000
	Positive	-13.26	5.74	-0.35	-2.31	0.025
	Negative	38.02	11.83	0.45	3.21	0.002
	Compound	6.74	2.01	0.58	3.36	0.001

Note. $p < 0.05$

Table 4에서 보듯, 긍정 에세이의 경우 총점 20점 중 평균은 10.92점으로 나타났다. 감성분석의 평균을 비교하면 긍정이 0.23, 부정이 0.05인데 반해 중립은 0.72로 나타났다. 하지만 통합 값을 기준으로 하면 0.78로 전체적으로는 긍정의 감성이 강하게 나타나고 있음을 확인할 수 있다. 본 연구가 찬성과 반대의 입장을 기술한 설득적 에세이 분석에 초점을 두고 있는 만큼 중립적 감성분석의 결과를 제외하고 긍정, 부정, 통합 값을 변인으로 하여 에세이 평가의 총점에 미치는 영향을 다중회귀분석을 통해 분석한 결과, Table 5와 같이 통합 값이 예측 변인으로 가장 큰 영향력을 보였다($\beta = 0.58, p = 0.001$). 물론 긍정과 부정의 감성도 유의한 수준에서 총점에 영향을 주는 것으로 나타났으나 한 가지 흥미로운 점은 찬성의 입장을 기술한 에세이 분석에서 긍정의 감성이 부적 영향을 미치고 있다는 점이다($\beta = -0.35, p = 0.025$). 이는 총점이 낮은 학습자일수록 단순히 성형수술의 장점을 나열하는 방식으로 긍정 어휘에 치우친 어휘 사용의 양상을 보였고 이것이 총점에 부정적인 영향을 준 것으로 보인다. 반면 상대적으로 총점이 높은 학습자들은 성형수술이 필요한 사람들이 가지는 아픔을 언급하거나 성형수술의 단점과 비교하면서 성형수술의 장점을 강조했기 때문에 긍정적인 어휘(e.g., confidence, helping, curing etc)뿐만 아니라 부정적인 어휘(e.g., scar, handicap, ugly etc)도 적절히 구사했고 이것이 총점에도 영향을 주었다고 판단된다. 결정계수인 R^2 은 0.21($Std. Error = 3.30$)로 나타나 이러한 감성관련 독립변수가 종속변수인 에세이 점수 변화의 21%를 설명한다고 할 수 있다. 또한 각 채점 영역(과제수행, 내용, 구성, 언어사용)별로 동일한 방식으로 다중회귀분석을 실시한 결과, 모두 동일 패턴의 결과가 나타났다. 결과적으로, 이러한 결과는 긍정 에세이의 감성분석의 경우 감성분석의 통합 값이 에세이 총점을 예측하는 도구로 활용될 수 있다는 것을 의미한다.

Sentiment Analysis as a Predictor for Scores of the Negative Essay

다음으로 부정 에세이에 대한 분석 결과는 아래의 Table 6과 Table 7에 제시된 바와 같다.

Table 6. Descriptive statistics of positive essays

Variables	M	SD	n
Total Score	10.65	4.04	60
Positive	0.11	0.05	60
Negative	0.12	0.06	60
Neutral	0.77	0.06	60
Compound	-0.05	0.67	60

Table 7. Multiple regression analysis on sentimental factors as a predictor for the positive essay score

Model	Predictor Variables	Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	β	t	Sig.
Positive Essay	(Constant)	11.55	1.97		5.87	0.000
	Positive	-5.45	16.73	-0.07	-0.33	0.746
	Negative	-2.16	16.15	-0.03	-0.13	0.894
	Compound	1.08	2.06	0.18	0.52	0.604

Note. $p < 0.05$

Word Cloud Analysis of the Persuasive Essay

Word cloud of Positive Essays ($n=60$)

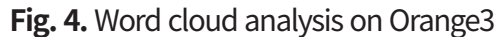


Table 8. Top 20 words of each 60 essays based on the two perspectives of 'positive' and 'negative'

223

공정과 부정 에세이에서 나타난 최상위 빈도 20개의 어휘를 비교해 보면 각 20개의 어휘 중 9개는 공통적으로 공정과 부정의 에세이에서 상위 빈도의 어휘로 나타났다. 긍정 에세이만 사용된 어휘는 성형수술이 필요한 사람들이 가진 문제점을 묘사하는 어휘(e.g., handicap, ugly)와 이를 극복하기 위한 성형수술의 장점을 기술하는 데 사용된 어휘들(e.g., curing, beautiful, advantages)로 구성되어 있다. 반면 부정 에세이에만 사용된 어휘는 긍정 에세이와는 달리 대부분 부정적인 결과를 기술하는 데 사용된 어휘(side effects, disadvantages, disagree, bad, cause)로만 구성되었다. 즉 이는 부정 에세이에서는 긍정 에세이와는 달리 성형수술의 장점을 제시하며 단점과 비교하는 방식보다는 단점 자체만 강조하여 기술하는 경향을 보였다는 것을 의미한다. 그리고 두 입장의 에세이에서 공통적으로 사용된 어휘는 주제와 관련하여 입장 차에 상관없이 사용할 수밖에 없는 어휘들(e.g., plastic surgery, people)이 주를 이루었다.

Conclusion and Implications

본 연구에서는 ‘Orange3’라는 파이썬(Python) 기반 텍스트 마이닝 프로그램을 통해 설득적 에세이의 감성 분석을 시도하였다. 그리고 이를 통해 감성분석이 설득적 에세이의 논조분석에 어느 정도 적용 가능한지를 가늠해 보고자 하였다.

본 연구에서는 먼저 Vader 방식의 감성분석(긍정, 부정, 중립, 통합)이 긍정과 부정 에세이의 총점에 미치는 영향을 살펴보았다. 그 결과, 긍정 에세이에서는 긍정, 부정, 통합의 감성분석 결과 값이 총점에 모두 유의한 수준에서 영향을 미쳤다. 특히 통합 값은 에세이의 전체 총점은 물론 과제수행, 내용, 구성, 언어사용의 4개 세부 채점영역 점수를 예측하는 변인으로서 가장 타당한 감성 요인이라는 것을 확인할 수 있었다. 반면 부정 에세이에서는 감성분석의 통합 값이 부정의 감성을 보였다는 것 외에는 모든 분석에서 유의한 결과를 보이지 못했다. 이는 긍정 또는 부정 중 하나의 논조를 선택하여 설득적 에세이를 작성하더라도 한쪽 입장을 일방적으로 주장하기보다는 양쪽의 입장을 비교하면서 한쪽의 입장을 강조하는 글 전개 방식을 채택할 경우 더 높은 점수를 기대할 수 있다는 것을 시사한다. 이와 같은 추정은 워드클라우드 분석을 통한 어휘 분포의 양상에서도 어느 정도 확인할 수 있었다.

본 연구의 시사점을 크게 두 가지로 요약하면 다음과 같다.

첫째는 긍정 에세이에서만 감성분석의 결과가 유의했지만 본 연구의 결과는 설득적 에세이 점수 예측을 위한 감성분석의 적용 가능성을 보여주었다는 점에서 의의가 있다는 점이다. 둘째는 장단점 중 하나의 입장을 취하여 설득적 에세이를 작성하더라도 장단점을 비교하며 주장하는 방식이 설득적 에세이 쓰기의 전략으로 보다 효과적이고 이를 체계적으로 지도할 필요가 있다는 점이다.

하지만 이러한 시사점에도 불구하고 본 연구에 사용된 데이터의 양이 크지 않았다는 점과 설득적 에세이의 논조에는 다양한 요인이 영향을 미칠 수 있는데 감성분석에만 초점을 두고 분석하였다는 점은 본 연구의 제한점이다. 따라서 추후 좀 더 방대한 데이터를 활용하여 다양한 요인을 분석하는 연구가 수행되기를 기대한다.

References

- Alm, A., & Nkomo, L. M. (2020). Chatbot experiences of informal language learners: A sentiment analysis. *International Journal of Computer-Assisted Language Learning and Teaching*, 10, 51-65.
- Arguedas, M., Xhafa, F., Casillas, L., Daradoumis, T., Peña, A., & Caballé, S. (2018). A model for providing emotion awareness and feedback using fuzzy logic in online learning. *Soft Computing*, 22, 963-977.
- Braddock, R. R., Lloyd-Jones, R., & Schoer, L. (1963). *Research in Written Composition*. Illinois, USA: National Council of Teachers of English.
- Demsar, J., Curk, T., Erjavec, A., Gorup, C., Hocevar, T., Milutinovic, M., Mozina, M. ... Zupan, B. (2013). Orange: Data mining toolbox in python. *Journal of Machine Learning Research*, 14, 2349-2353.
- Farra, N., Somasundaran, S., & Burstein, J. (2015). Scoring Persuasive Essays Using Opinions and Their Targets. In J. Tetreault, J. Burstein & C. Leacock (Eds.), *Proceedings of the 10th Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 64–74). Denver, Colorado: Association for Computational Linguistics.
- Fowler, J. K. (2019). *Enhancing Writer's Voice in Argumentative and Expository Essay Writing Using the Embedded Voice Technique* (Doctoral dissertation). University of South Carolina, Columbia, USA.
- Gentry, B. (2010). *Improving Student Writing in The Political Science Classroom: Strategies for Developing Student Voice and Originality*. Paper presented at the Southern Political Science Association Conference, Atlanta, GA.
- Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. Paper presented at the 8th International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI.
- Jhang, S., Lee, K., Park, H., Song, W., Jung, H., Lee, S., & Kim, J. (2019). Sentiment analysis of Shakespeare's tragedy plays and their major male and female characters through dialogue corpora. *Journal of Language Sciences*, 26, 115-147.
- Lin, H.C. K., Chen, N.S., Sun, R.T., & Tsai, I.H. (2014). Usability of affective interfaces for a digital arts tutoring system. *Behaviour & Information Technology*, 33, 105-116.
- Liu, B., Hu, M., & Cheng, J. (2005). Opinion Observer: Analyzing and Comparing Opinions on the Web. In A. Ellis & T. Hagino (Eds.), *Proceedings of the 14th International Conference on World Wide Web* (pp. 342-351). Stroudsburg: Association for Computational Linguistics.
- Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior*, 31, 527-541.
- Shin, D. (2020). A validity study on the item type of 'identifying the emotional state or mood' in the English section of CSAT based on sentiment analysis. *Journal of the Korea English Education Society*, 19, 159-179.
- Shin, D., Kim, I., Park, J., & Kim, Y. (2016). *Searching for Ways to Revitalize School Education for Narrative and Essay-type Paper-based Assessment* (Gwangju Metropolitan Office of Education Tech. Rep. No. 2016-54). Jincheon: Korean Educational Development Institute.
- Stab, C., & Gurevych, I. (2017). Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43, 619-659.
- van Eemeren, F. H., Grootendorst, R., & Henkemans, F. S. (1996). *Fundamentals of Argumentation Theory: A Handbook of Historical Backgrounds and Contemporary Developments*. New York, NY: Routledge, Taylor & Francis Group.

- Wilson, T. A. (2008). Fine-grained Subjectivity and Sentiment Analysis: Recognizing the Intensity, Polarity, and Attitudes of Private States(Doctoral dissertation). University of Pittsburgh, PA, USA.
- Zhou, J., & Ye, J. M. (2020). Sentiment analysis in education research: A review of journal publications. Interactive Learning Environments. Retrieved from <https://doi.org/10.1080/10494820.2020.1826985>

Author Information

Shin, Dongkwang: Gwangju National University of Education, Professor, First Author

ORCID: <https://orcid.org/0000-0002-5583-0189>