# [ 표 지 ]

**논문 준비 계획서**

**코로나19사태와 제21대 대한민국 총선보도의 프레임 분석:**

**빅데이터 및 단어임베딩, 감성분석을 이용하여**

**Analysis of South Korean Journalism Framing on COVID-19 and 21st South Korea General Election:**

**Big Data, Word Embeddings, and Sentiment Analysis Leveraging**

지도교수 : 이경숙 교수님

제출일 : 2020년 6월 5일

고려사이버대학교 융합정보대학원

정보기술 전공

정 훈 영

학번: 201950009

# [ 목 차 ]

[[ 표 지 ] 1](#_Toc43701323)

[[ 목 차 ] 2](#_Toc43701324)

[[ 요 약 ] 3](#_Toc43701325)

[[ 본 문 ] 4](#_Toc43701326)

[I. 논문의 주제 4](#_Toc43701327)

[II. 서론 4](#_Toc43701329)

[1. 주요선행연구 5](#_Toc43701330)

[2. 문제의 제기 6](#_Toc43701331)

[3. 연구의 목적과 필요성 8](#_Toc43701332)

[4. 연구문제 8](#_Toc43701333)

[III. 이론적 배경 9](#_Toc43701334)

[1. 언론의 정치적 프레임워크(Framework)와 신뢰성 9](#_Toc43701335)

[2. 프레임(Frame) 분석의 목적과 방법 11](#_Toc43701336)

[3. 워드임베딩(Word Embedding)기법과 감성분석(Sentiment Analysis) 13](#_Toc43701337)

[IV. 연구방법 14](#_Toc43701338)

[1. 연구대상 14](#_Toc43701339)

[2. 연구도구 14](#_Toc43701340)

[3. 연구절차 및 자료수집 15](#_Toc43701341)

[4. 자료분석방법 16](#_Toc43701342)

[5. 연구설계 17](#_Toc43701343)

[V. 연구결과 – 예상되는 연구결과 18](#_Toc43701344)

[[ 부 록 ] 18](#_Toc43701345)

[VI. 참고문헌 18](#_Toc43701346)

# [ 요 약 ]

한국언론은 각각의 고유한 정치 이념적 편향성을 가지고 있고, 이는 언론이 생산하고, 유통시키는 뉴스콘텐츠에서 프레임형식으로 발현된다. 정치적 이념의 가장 큰 특질은 설득이다. 이념의 주체에 있어서, 정치적 이념은 지켜야 할 본질이며, 프레임은 이를 관철시키는 도구이다. 반면, 언론이 지켜야 할 가치는 진실이다. 저널리즘의 기본적인 기능은 객관적인 사실을 바탕으로 진실을 전달하는 것에 있다. 따라서, 언론이 정치 커뮤니케이션 매개자의 역할을 넘어, 주체적으로 정치적 이념을 관철시키려 할 때, 이러한 과잉된 프레이밍으로 편향성을 드러낸다. 저널리즘의 ‘진실’이란, 보편타당한 진실과는 다르다. 하지만, 진실의 본질을 전달하는 것과 프레임으로 포장된 진실을 전달하는 것은 다르다. 이용자에게 있어서 언론이 생산하고 유통하는 뉴스콘텐츠는 세상을 바라보는 ‘창(window)’이다. 따라서, 언론이 객관적사실에 기반한 공정한 진실을 전달하지 않는다면, 이로 인해 사회적 혼란이 가중되고, 극단적으로는 우리사회를 지탱하는 민주주의의 붕괴로 이어질 수도 있다. 한국언론의 특별한 정치 이념적 편향성은 보수와 진보, 또는 우파와 좌파의 양진영으로 쉽게도 분열되고, 정권의 교체와 더불어 권력투쟁의 양상으로까지 비춰지고 있다. 따라서, 본 연구에서는 이러한 언론의 정치 이념적 편향성이 정치적인 사안을 넘어, 사회적사안 또는 의료적인 사안까지 영향을 미치는지 연구하고, 그 실태를 검증해보고자 한다. 구체적으로는, 대표적인 정치적 사안인, ‘제21대 대한민국 총선보도’에 가장 큰 영향을 끼친 프레임은 무엇인가 살펴 보고, 이러한 프레임은 정치 이념적 편향성을 가지고 있는지 가늠해 본다. 또한, ‘제21대 대한민국 총선보도’에서의 프레임이 ‘코로나사태’에도 동일하게 주요 프레임으로 동작하였는지, 또는 부차적으로 사용되었는지 알아보고자 한다. 이를 통해 정치 이념적 편향성이 언론기관전체의 편향성을 대별(大別) 할 수 있는지, 또는 이용자의 선입견을 넘어 공정한 언론관을 유지하고 있는지 알 수 있을 것이다.

선행연구와의 공통점과 차이점은, 기존연구의 이야기(narrative)구조 분석방식과 지침 기반의 전일적 접근(manual holistic approach)방식의 일부를 채용한다. 이야기구조는 ①국정운영의 ‘위험성’을 강조 ②국정위험의 ‘원인’을 규정, ③ ‘대책’으로 국정안정을 지지/정권심판, 하는 주제적 특성으로 갖게 될 것을 전제한다. 차이점으로는 주제모델(Topic Model)의 의미해석(Semantic Analysis) 알고리즘(algorithm)의 단점을 보완하기위해, 기계학습(Deep Learning)의 일종인 워드임베딩(Word Embedding)을 사용한다. 1차(제21대 대한민국 총선보도)는 및 2차(코로나 사태)의 뉴스콘텐츠들은 심층학습을 통해 데이터모델로 구축된다. 1차와 2차에서 도출된 데이터모델은 유사도분석을 시행하여 유사성을 추출한다. 1차와 2차에서 도출된 데이터모델이 ‘여당 우호적’인 이야기구조와 부합여부를 검증한다. 이를 통해 뉴스콘텐츠의 프레임의 정치 이념적 편향성 여부를 가늠할 것이다.

# [ 본 문 ]

## **논문의 주제**

코로나19사태와 제21대 대한민국 총선보도의 프레임 분석:

빅데이터 및 단어임베딩, 감성분석을 이용하여

Analysis of South Korean Journalism Framing on COVID-19 and 21st South Korea General Election:

Big Data, Word Embeddings, and Sentiment Analysis Leveraging



## **서론**

2020년 4월15일, 제21대 대한민국 총선이 시행되었다. 코로나 바이러스가 전세계적으로 대유행중인 시점에 세계에서 시행된 전국단위의 유일한 선거였다. 최종 투표율은 66.2%로써, 제20대 대한민국 총선의 58.0% 보다 8.2%높은 결과였다1). 보통 선거에는 정치적 대립(Political Conflict)구도가 존재하며, 이를 실체적으로 드러냈을 때, 프레임전쟁이라고 한다. 많은 언론들은 ‘코로나19’가 ‘제21대 대한민국 총선’의 프레임전쟁을 약화시켰다고 진단한다. 그렇다면, ‘코로나19사태’가 ‘제21대 대한민국 총선’의 프레임인가? 프레임은 설득력을 동반하여야 한다. 따라서, ‘코로나19사태’ 그 자체를 프레임이라고 할 수는 없다. 제21대 총선의 프레임은 ‘국정안정’ 이었다2). 언론이 생산하고 유통하는 뉴스는 항상 ‘진실’이어야 한다. 하지만, ‘진실’은 여러가지 측면을 가질 수 있고, 진실의 한 단면을 강조하여, 이에 대한 강제적 수용을 이용자에게 요구하는 것을 프레이밍(framing)이라고 할 수 있다. 한국 언론에서 가장 많이 사용되는 프레임의 차원은 주로 ‘갈등’이었다(이희영·김정기,2016). 갈등은 이해관계의 충돌에서 비롯된다. 이해관계의 충돌은 이해관계자간의 이익관계가 상충될 때 발생한다. 한국사회가 여타 사회집단보다 이기적이라고 할 만한 증거는 없다. 예측되는 주요한 갈등의 요인은 ‘정치적 이념’이다. 한국의 정치적 이념은 서양의 정치적인 이념과는 달리, 깊은 토론과 오랜 논쟁을 통해 발전하지 못하였다. 우리의 정치이념은 급속한 근대화사회로의 발전과정에서 그 형식과 근간을 구축했으며, 공산주의에 대한 민주주의의 대항적 편향성위에서 발전하였다. 한국의 이념적 정치성향이란, ‘보수’ 또는 ’진보’를 논하기에 앞서, 그 이념적 근간이 정치체제 및 경제시스템과 맞닿아 있어, 이념의 붕괴란, 곧 정치체제와 경제시스템의 붕괴를 의미하는 형상이 되어버렸다 (김병곤,2011). 따라서, 언론의 이념적 담론에 있어서, 보수언론은 주로 기득권층의 논리를 대변하고, 진보언론은 이에 대한 적대적 논리로 반박하는, 사회집단간의 대리전양상으로 치닫고 있다. 따라서, 본 연구에서는 먼저, 한국 언론의 정치 이념적 편향성을 가장 큰 프레임워크(Framework)로 예측하고 있으며, 도출된 상위의 프레임(Frame)이 이에 귀속될 수 있는지에 대한 논의를 살펴 본다. 또한, 기존의 뉴스 프레임 분석 연구에서 지지할부분과 한계점을 파악하여, 이에 대한 개선된프레임분석방법을 제시해보고자 한다

1. 중앙선거관리위원회(2016. 4. 13. 실시 제20대 국회의원선거 투표율 분석)
2. 중앙선거관리위원회 제21대 국회의원선거 ‘현안’ 조사결과

### **주요선행연구**

* 한국 보수주의의 이념적 특징.

김병곤(2011). 역사비평, 8-35.

* 뉴스 틀 짓기 연구의 두 개의 뿔.

이준웅(2009). 커뮤니케이션 이론, 51,123-166.

* 미세먼지 재해 보도의 프레임 분석.

이준웅·김성희(2018). 한국언론학보, 62(4), 125-158.

* 한국 정파언론 환경의 특수성은 보수와 진보 수용자의 매체 태도와 이용에 차별적 영향을 미치는가?

현기득·서미혜(2019). 한국언론학보, 63(2), 46-76.

* 질적 메타분석을 통한 뉴스프레임의 유형.

이희영·김정기(2016). 한국언론학보, 60(4), 7-38.

* 보도의 신뢰 하락 원인 분석과 책임 저널리즘을 통한 방송 보도의 신뢰성 확보 방안 연구. 김세은·심훈·박영흠(2018). 방통융합정책연구, KCC-2018-39.
* 사실 확인인가, 사실 만들기인가.

김창숙(2019).이화여자대학교 대학원,박사학위논문.

* 알고리즘 기반의 프레임 추출로 본 정부와 언론의 원자력 프레임 경쟁 양상 : 프레임 자동 추출 기법의 효율성과 타당성 연구.

임종섭·김진희(2020). 언론과학연구, 20-1, 151-190.

* Framing: Toward Clarification of a Fractured Paradigm,

Robert M. Entman, Journal of Communication, Volume 43, Issue 4, December 1993, Pages 51–58

* Efficient estimation of word representations in vector space.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). arXiv preprint arXiv:1301.3781.

* Distributed representations of words and phrases and their compositionality

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). In Advances in neural information processing systems (pp. 3111-3119).

* A neural probabilistic language model.

Bengio, Y., Ducharme, R., Vincent, P., & Jauvin, C. (2003). Journal of machine learning research, 3(Feb), 1137-1155.

* Learning word vectors for sentiment analysis.

Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts. 2011. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1 (HLT ’11). Association for Computational Linguistics, USA, 142–150.

### **문제의 제기**

#### **진실과 프레이밍(Framing)**

언론이 이념을 가지고 있는 것은 문제가 될 수 없다. 언론은 정치적이다. 한국 언론의 정치 이념적 편향성의 여부또한 이미 기정 사실화 되고 있다. 언론은 권력집단이다(남시욱,1997). 언론은 시민사회와의 암묵적인 계약을 통해 이러한 권력을 위임받았다고 할 수 있다(김세은· 심훈·박영흠,2018). 역사적으로 언론의 첫번째 의무는 진실을 전하는 것이었다. 일반적으로 통용되는 보편타당한 진실과 저널리즘의 진실의 의미에는 차이가 있다. 저널리즘이 추구하는 진실은 행위의 기준이 되는 실질적이고, 기능적인 진실이다(김창숙,2019). 따라서, 언론의 위임받은 권력은 진실을 전달하는 행위를 위하여 행사되어야 한다. 만약, 언론이 주체적으로 정치적 이념을 설파하거나, 특정 주체의 주장을 대변하려 한다면, 사회계약은 해체되고, 취재와 보도 행위의 정당성이 취약해지며, 사회적 신뢰는 붕괴된다. 따라서, 이러한 정보와 언론의 신뢰체계 붕괴는 민주주의 정치 시스템의 약화에 직접적인 위협으로 작용하게 된다. 민영(2016)은 언론이 정치 체계 구조와 유기적인 관계이며, 한국 언론의 신뢰와 정치 신뢰가 정적인 상관관계에 있음을 규명하였다. 정치집단의 정치 행위에 대한 정보는 언론의 설득력의 여하에 따라, 이용자의 수용자기제에 부합하게 되는 기대효과를 가지게 되므로, 언론 권력과 정치 권력은 상호 유착관계를 형성하기 쉽게 된다. 정치 권력과 이에 호응하여 권력적 지위를 유지하는 언론 권력, 이것은 언론의 정치 이념적 편향성을 가중시키는 중요한 요인이다. 그렇다면, 한국 언론은 다만 기득권의 유지를 위해 정치 이념적 편향성을 가지게 되는가? 한국 언론은 근대화를 거치면서 고유의 정치 이념을 계승하게 되었다. 한편, 이에 반하는 정치적 이념또한 대두되었다. 정치적 이념의 가장 큰 특질은 설득이다. 정책학 연구자들은 프레임을 사회 행위자들이 의미를 만들기 위해 동원하는 설득 장치로 이해하며, 이 행위자들이 프레임을 내세워 정책 문제의 정의를 둘러싼 상징적 힘겨루기를 한다고 설명한다(임종섭·김진희,2020). 즉, 정치적 이념은 가치이며, 프레임은 도구이다. 한국 언론은 해외언론에 비해, 특수한 환경에 있어, 정치 이념적 편향성을 보수와 진보의 구도로 쉽게 구분할 수 있다(현기득·서미혜, 2019). 사회는 정반합(thesis, antithesis, synthesis)의 대립구도에서 발전한다3). 보수와 진보의 대립관계를 사회의 분열이라고 진단하기는 어렵다. 보수가 말하는 진실(thesis), 진보가 말하는 진실(antithesis), 이에 대한 합의(synthesis)의 결과로서, 사회는 발전한다. 하지만, 대립관계를 넘어 적대적인 관계라고 표현할 수 있을 때 사회적 분열과 정치제제의 붕괴가 발생한다.

1. And even when a society has got upon the right track for the discovery of the natural laws of its movement – and it is the ultimate aim of this work, to lay bare the economic law of motion of modern society – it can neither clear by bold leaps nor remove by legal enactments, the obstacles offered by the successive phases of its normal development. But it can shorten and lessen the birth-pangs. 카를 마르크스는 자본주의의 이러한 모순과 대립을 통한 사회발전을 자연스러운 현상으로 인식하고 있었다. 실질적으로, 카를 마르크스가 현대의 자본주의의 토대와 근간을 마련하였고, 사회과학을 과학적방법으로 과학화했다는 것은 주지의 사실이다. (Capital: Preface to the First German Edition (Marx, 1867): The Process of Production of Capital)

한국의 특수한 정치환경과 이념적 정치성향이란, ‘보수’ 또는 ’진보’를 논하기에 앞서, 그 이념적 근간이 정치체제 및 경제시스템과 맞닿아 있다. 따라서, 이념의 붕괴란, 곧 정치체제와 경제시스템의 붕괴를 의미하는 형상이 되어버렸다. 이준웅(2017)은 ‘가짜뉴스’문제의 기인으로, 프래임경쟁을 지목하였다. 본 연구에서는 ‘가짜뉴스’의 문제를 언론에 대한 신뢰성결여의 실체적인 현상으로 이해하고 있다. 따라서, 언론이 진실을 이용자에게 공정하게 전달할 수 있다면, 이는 신뢰성의 회복으로 이어지며, 자연스럽게 잦아들 현상이라고 할 수 있다(민영, 2016). 이에, 본 연구에서는 한국 언론의 정치 이념적 편향성이 과잉된 프레임전쟁의 양상으로 정치적 사안을 넘어, 사회전반의 뉴스콘텐츠에 영향을 주고 있고, 이러한 현상이 더욱 고착화 되고 있는지, 또한, 이러한 정치 이념적 편향성의 정도는 어떻한지 기계적인 해석이 필요하다는 문제의식에서 시작한다.

#### **프레임(Frame) 분석과 추출방법**

프레이밍(framing)은 상징적인 단어나 문화적으로 친숙한 기호의 반복적인 사용을 통하여, 주로 이용자의 인지적 편향성에 호응하는 효과를 발휘한다(Robert M. Entman, 1993). 접근방법은 대별(大別)적으로 연역적인 접근방법과 귀납적인 접근방법으로 구분 할 수 있다. 연역적 접근은 선행 연구의 프레임이 텍스트에 나타나는 정도를 파악한다. 이 접근은 새로운 프레임을 확인할 수 없고 일부 프레임을 생략할 수 있다. 귀납적 접근은 텍스트에서 프레임을 발견하는데, 새로운 프레임을 찾을 수 있으나 시간과 노력이 필요하고 대용량 데이터를 분석하기 어렵다. 또한, 기본적으로 각각의 접근방법론들은 각각의 장단점을 가지고 있어 사안에 따라 접근방법을 달리하는 경우가 많다. 한편, 알고리즘과 구조모형을 이용한 새로운 형태의 분석방법도 등장하였다. 이준웅, 김성희(2018)는 구조적 주제모형(Structural Topic Modeling)을 이용하여, 특정한 사안에 대한 잠재적 의미 분석을 알고리즘을 통하여 시도하였으며, 임종섭 등(2020)은 의미 연결망(Semantic Network)과 군집도(corpus)분석을 통해 프레임추출을 시도 하였다. 두 연구의 공통점은 데이터마이닝(Data Mining) 및 자연어처리(NLP)에서 많이 사용되는 주제모델(Topic Model)이 기반이라는 점이다. 주제모델은 문장안에서 사용되는 단어의 빈도수와 밀접한 관계가 있다. 의미연결망또한 문서 단어 행렬(Document-Term Matrix)의 관계형성 알고리즘(Term Frequency-Inverse Document Frequency)으로써, 데이터마이닝(Data Mining)기법이라고 할 수 있다. 의미연결망은 최근 들어 빅데이터분석에 빈번하게 사용되는 기법으로, 시각화 하기 쉽고, 프레임의 특성을 고려했을 때 유용한 기법이라고 할 수 있다. 이 모형들의 단점은 ‘단어의 빈도수’에 의존적관계가, ‘프레임의 의미의 지향성’을 제시하기 어렵다는 점이다. 프레임은 본질적으로 ‘선택’의 맥락에서 해석되어야 한다. 특정 단어의 사용빈도수를 의미연결망에서 보여주는 상위의 결과는 ‘주제’에 대한 추정이 아닌, ‘수치’상의 결과로써, 언어의 화용론적 해석은 어렵다고 할 수 있다. 따라서, 사안이 특정되어 있지 않거나, 내용을 파악하고 있지 않다면, 결과에 대한 해석은 어려워 지고, 추출결과에 대한 인위적인 조작이 필요하다. 본 연구에서는 이러한 단점을 보완하기 위해 워드임베딩(Word Embedding)을 사용하고자 한다.

### **연구의 목적과 필요성**

범주화, 프레임은 범주화된 단편적인 ‘진실’이다. 사회에는 다양한 ‘진실’이 공존하여야 한다. 특정 집단이 추구하는 진실만이 우위를 구가(謳歌)하는 사회는 더이상 평등하고, 공정한 사회라고 부를수 없다. 뉴스콘텐츠또한 다양한 프레임으로 구성되어야 한다. 극단적인 소수의 프레임이 다수의 뉴스콘텐츠를 지배한다면, 언론은 더이상 진실을 보도하지 않는다고 할 수 있고, 대중의 신뢰는 멀어질 것이다. 하지만, 제기된 문제점들을 살펴 보았을 때, 한국 언론의 정지척 편향성에 대한 거대담론은 보수와 진보집단간의 ‘공격 저널리즘’의 악순환을 되풀이 하는 것으로 보여지고 있으며, 정치 이념적 프레임이 정치적 사안을 벗어나, 사회, 문화, 경제전반에 걸쳐 지배적인 프레임으로 비춰진다(최영재,2011;현기득·서미혜,2019). 언론의 정치적 특질을 부정하지는 않는다. 다만, 과잉된 정치 프레임전쟁의 양상이 정치적 사안을 넘어, 사회전반의 뉴스콘텐츠에 영향력을 행사할 때, 진영논리로 비화하는 것은 경계해야 한다. 따라서, 본 연구에서는 이러한 한국 언론의 실태를 보다 기계적으로 검증할 수 있다는것에서 그 필요성이 있다고 할 수 있다. 또한, 본 연구의 기본적인 목적은 기계적인 프레임분석방법을 통하여, 보다 면밀하게 프레임을 분석하는 것이며, 부차적으로는 추출된 각 프레임의 기계적인 해석을 통하여, 상호 비교하는 것이다.

### **연구문제**

1. 각 언론사별 생산하는 ‘제21대 대한민국 총선’뉴스와 ‘코로나19사태’뉴스에서 추출한 프레임의 주제유목(categorize)은 ‘원인’, ‘결과’, ’대책’으로 구분할 수 있는가?
2. 각 언론사별 생산하는 ‘제21대 대한민국 총선’뉴스와 ‘코로나19사태’뉴스에서 추출한 프레임의 주제유목은 서로 유사한가?
3. 각 언론사별 생산하는 ‘제21대 대한민국 총선’뉴스와 ‘코로나19사태’뉴스에서 추출한 프레임의 주제유목은 아래의 이야기(narrative) 구조 주제유목과 유사한가?
   1. 내용: 코로나및 경제위기등에 대한 위험성을 강조
   2. 내용: ‘원인’으로 대통령 및 정권여당의 실정(失政)을 강조, 또는 야당의 국정발목잡기를 강조.
   3. 내용: ‘대책’ 및 ‘결과’로써, 국정안정을 위해 여당지지 또는 야당지지

## **이론적 배경**

### **언론의 정치적 프레임워크(Framework)와 신뢰성**

일반 대중이 원하는 진실이란, 보편타당한 실존적인 진실이다4). 저널리즘이 추구하는 진실이란, 공리적이며, 기록되고 확인될 수 있는 사실들이고, 확인에 의해 변화할 수 있는 진실들이다(김창숙,2019). 저널리즘에서의 진실이란, 사실을 기반으로 하되, 현재상황에서 얻을 수 있는 최선의 객관적 선택이며, 언제든지 변화할 수 있는 입증가능한 정보를 의미한다. 동일하게, 프레이밍(framing)또한 본질적으로 사안의 특질을 구분하는 선택을 동반한다(Entman, 1993). 이희영, 김정기(2016)는 질적 메타분석연구를 통하여, 한국 언론의 뉴스의 포괄적인 프레임을 구분하였다. 이를 공리화할 수 없다는 한계점에도 불구하고, 포괄적인 한국 뉴스의 프레임특성의 일반화의 시도라는 점은 주목할 만하다. 동 연구에 의하면, 한국 언론에서 부각 되는 주프레임은 ‘갈등’이었으며, 이에 대한 ‘원인’, ‘결과’, ’대책’에 상당히 집중하고 있었고, 해당 프레임의 범주는 ‘정책’과 ‘민주합의’였다. 따라서, 한국 언론이 집중하는 이야기적인 프레임워크(narrative framework)는 ‘정치적’이라고 할 수 있고, 한국 정치의 특질상 보수-진보, 우익-좌익의 이분법적 이념논쟁이 부각 될 수 밖에 없다. 김병곤(2011)은 한국언론이 보수주의를 중심으로, 진보와의 대립관계에 서게된 배경을 아래와 같이 진술한다.

“군부에 의한 권위주의시대와는 다르게 보수세력의 이념적 정당성을 주장하고 유지함으로써

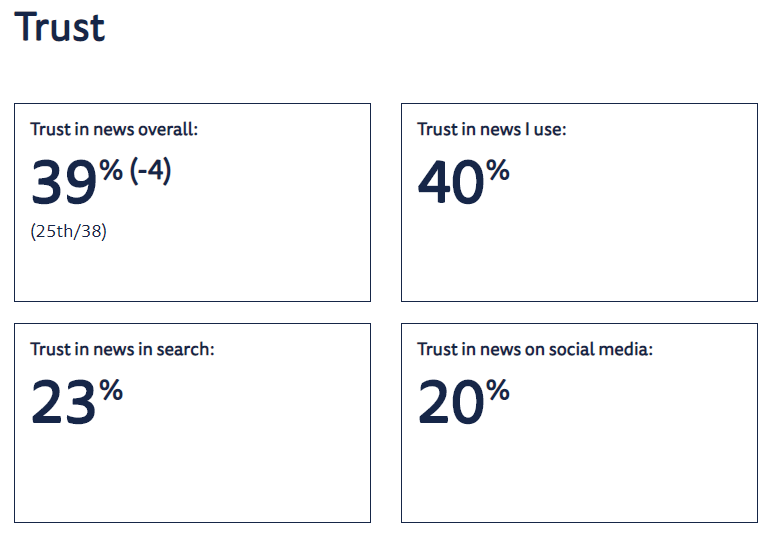
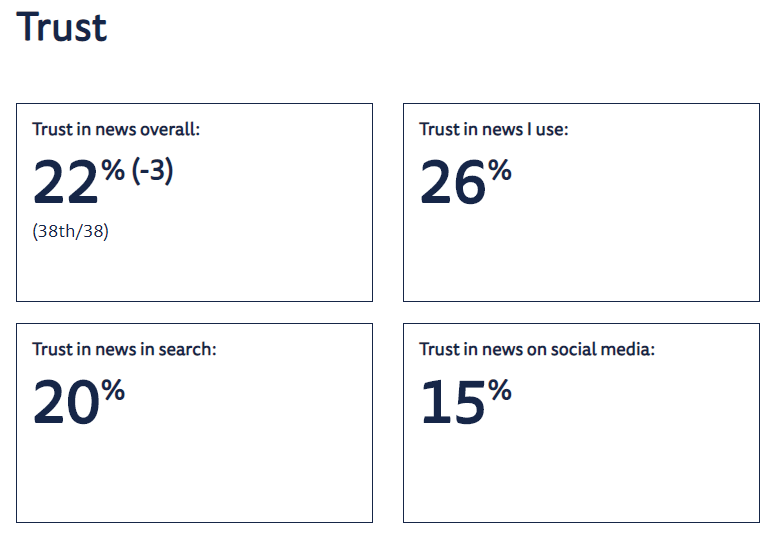
기득권층의 이익을 보호하는 역할은 국가가 아니라, 이른바, 조·중·동으로대표되는 보수언론이 떠맡게 되었다. 보수언론을 중심으로 종교와 보수단체 등이 전면에 나서기 시작하면서 보수세력의 결집은 한층 확대되었다. 보수정권이 담당하던 반공주의는 이제 보수언론과 보수단체들의 중심 가치로 주장되었다.”

사회집단이 이미 이분화되어 언론의 뉴스를 편향적으로 수용하고 있다는 것도 주지의 사실이지만, 한국 언론이 정권의 몰락과 득세에 따라 ‘편향 전략(bias strategy)’을 사용하며, 극도로 편향된 뉴스 프레이밍으로 상대진영을 공격하는 극단적 공격 저널리즘(attack journalism)이 이에 대한 악순환을 부추기는 것도 사실이다(현기득·서미혜,2019; 엄현식·강재원, 2019; 최영재,2011). 프레임추출의 어려움은 이러한 뉴스또한 형식적으로는 객관주의로 은폐되어 있다는 것이다. 형식적 객관주의는 인용, 데이터의 조작적 정의등을 포함하여, 상대진영의 사소한 실수나 실언등의 인과관계를 침소봉대하는 등, 과잉된 사실을 진실로 포장하는 속성이 있다. 김세은 등(2018)의 연구에서도 언론의 신뢰성을 회복시키기 위해서 “큰따옴표 저널리즘”의 지양을 주문한다. 즉, 한국 언론 전반에 대한 신뢰는 정치 신뢰와 밀접한 관계에 있다고 할 수 있다. 언론은 정치 정보를 선택적으로 제공함으로써 ‘정치 제도’에 대한 특정한 기대를 창출하는데, 언론이 사회 구성원들의 기대에 부합하지 않아 신뢰를 잃을 때 정치 행위자들에 대한 평가와 신뢰 역시 하락할 수 있다는 의미이다(민영,2019). 그렇다면, 언론의 보도는 내용과 그 형식에 있어서, 모두 객관적 사실이어야 하는가?

1. 카를 야스퍼스에 의하면, 보편타당성이란, 각 개인간의 사유와 행위에서 관계를 통하여 합의된 도덕률에 가깝다.

<그림.1>은 2019년도 Reuters Institute Digital News Report의 한국언론에 대한 신뢰도 결과이다5). 조사대상국중 신뢰도 지수는 최하위이다. 한편, 상대적으로 한국보다 언론의 자유지수가 낮은 일본의 경우를 살펴보았다<그림.2>. 국경없는기자회(RSF)의 World Press Freedom Index 2020의 조사6)에 의하면, 일본의 언론자유도는 한국보다 많이 뒤쳐져있는 상황에서도, 언론에 대한 신뢰도는 높은 것으로 평가되었다. 이에 대한 시사점은 일본 언론을 본받아야 한다는 점이 아니다. 일본의 대중들은 언론의 자유, 진실등에는 관심이 없다. 鶴田達成(2018)의 연구에 의하면, 일본의 언론은 이미 정부와의 유착관계로 인해 민주주의 언론으로서의 기능을 상실했다고 진단한다. 일본 정치상황의 특수성은, 보수우익의 단일한 정치집단(Liberal Democratic Party)이 2차대전후부터 장기집권중이며, 극단적 우익성향은 더욱 심화되고 있다. 또한, 한국 언론에서는 쟁점화 되지 않는 정부언론(government speech), 정부가 언론의 자유를 제어하여, 국민들의 생각을 직접 통제하려는 행위에 대한 문제도 심각하다고 진단하고 있다(稲積重幸,2011). 김하나(2017)는 일본 언론의 보편적인 편집이념 ’不偏不党’이 ‘확인된 사실’ 또는 ‘정부의 공식적인 입장’만을 전달하고, ‘진실’을 전달하지 않는 것이 일본언론의 실태라고 말해주고 있어, 일본 언론은 정권의 대변인이며, <그림.2>의 수치는 그저 형식적인 숫자에 불과하다라는 것도 이해해야 할 것이다. 따라서, 한국 언론이 일본의 언론상황을 통해 얻어야 할 시사점은, 언론이 제4의 권력으로서, 공정한 역할을 수행하지 못하였을 때, 한국 언론또한 언론의 자유와 권리를 상실할 수 있으며, 국민들은 이를 자연스럽게 수용할 것이라는 점이다. 비트겐슈타인(Ludwig Josef Johann Wittgenstein)은 논고의 마지막을 이렇게 기술했다.

<그림.1>



<그림.2>

“7 말할 수 없는 것에 관해서는 침묵해야 한다.”

1. Reuters Institute Digital News Report 2019는 아래와 같이 한국언론의 신뢰도와 특징을 진단하였다. ‘개별적인 언론의 신뢰도는 높지만, 대표적인 신문브랜드 신문에대한 신뢰도는 낮다. 또한, 일본의 경우에는 사실성위주의 보도에 집중하는 공영방송인 NHK이외의 기타 브랜드는 흥미위주의 기사를 주로 제공하고 있어, 실질적으로는 공공성이 높은 정보가 대중에게 잘 전달되고 있지 않다고 생각할 수 있다. 따라서, 일본국민의 정치, 사회에 대한 무관심을 생각 할 수 있다.
2. 국경없는기자회(RSF)의 World Press Freedom Index 2020의 조사에 의하면, 한국은 23.70으로 42위, 일본은 28.86으로 66위이다.

그는 논고를 통하여 ‘말’의 ‘무게’를 진술하고 있다. 한국언론이 진실을 추구하여야 하는 이유는, 진실을 ‘말’할 수 있고, 이를 전달할 수 있는 권리와 책임, 그리고 사명이 있기때문이다. 또한, 많은 선행연구들에서 부각되는 단어가 있었다. ‘탈진실(post-truth)의 시대’. 탈 진실이란, 진실이 다른 고려사항에 비해 상대적으로 낮은 가치를 부여받고, 그 결과 진실에서 벗어나는 것에 대해

대중이 수용적인 경향을 보이는 현상을 의미한다. 윤성옥(2018), 이미나, 구교태(2019), 염정윤, 정세훈(2019)등의 연구에서 이용자의 ‘미디어 리터러시’강화를 주문하는 것은 이러한 탈진실의 시대에서 일반 대중들이 진실을 가늠할 수 있어야한다는 의미로 해석할 수 있다. 따라서, 한국 언론은 이러한 탈진실의 세태에서도, 과잉된 프레임전쟁을 넘어, 국민들이 신뢰할 수 있는, 진실을 가늠하는 척도이어야 할 것이다.

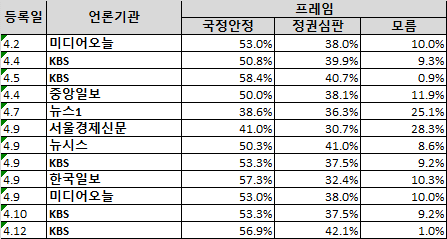
### **프레임(Frame) 분석의 목적과 방법**

프레임(frame)은 지각할 수 있는 몇가지의 단면중 한가지 특질을 선택한다. 프레임(frame)은 사안의 정의, 인과관계, 도덕적인 평가 또는 개선방법등을 설명하기 위한 수단이라고 할 수 있다(Entman, 1993; Gamson & Modigliani, 1989; Gitlin, 1980). 따라서, 뉴스콘텐츠에서 프레이밍을 사용하는 원인은 비교적 명확해진다. 이용자의 수용기제에 대한 커뮤니케이션 효과라고 할 수 있다(Entman, 1993). 그렇다면, 뉴스콘텐츠의 프레임을 분석하는 목적은 무었일까? 프레임의 특질이 선택이라면, 프레이밍의 목적은 설득이다. 설득은 동일하게 ‘정치적 이념’의 가장 큰 특질로써, 프레임의 실체를 언론 및 여타 사회권력, 정치권력과의 헤게모니(hegemony)쟁탈전이라고 규정하여도 무리가 없다(Carragee, K. M., & Roefs, W, 2004). 따라서, 뉴스프레임에 대한 분석은 주로 ‘누가 헤게모니를 주도하는가, 어떠한 수용기제가 작동하였는가’를 파악하기 위한 목적으로 연구된다. 뉴스 프레임분석은 통상적으로 질적연구의 영역에 해당하며, 구체적으로는 아래와 같은 접근방법들을 고려 할 수 있다(임종섭·김진희, 2020).

1. 해석학적 접근(hermeneutic approach), 사회 문화의 배경을 고려해 텍스트를 해석하고 설명함으로써 프레임을 파악한다. 하지만 주관적 분석에 의존해 신뢰성이 불투명하고, 프레임의 추출 과정이 명확하지 않으며, 연구자가 원하는 프레임만 발견하는 프레임의 선택 편중성이 발생할 수 있다.
2. 언어 접근(linguistic approach), 특정 단어가 프레임을 구성한다고 전제하고 특정 단어와 문장의 선택, 배치, 구조 등을 조사해 프레임을 식별한다. 그러나, 대규모 자료를 분석하기 어렵고 프레임 구분이 명확하지 않은 경우가 있다.
3. 지침 기반의 전일적 접근(manual holistic approach), 사전 연구에서 개발한 코드 지침으로 텍스트를 분석한다. 지침의 신뢰성과 타당성이 프레임 추출의 신뢰도와 연결되며, 연구자가 찾는 프레임만 파악하는 위험이 있고 인간에 의존해 주관성의 문제가 있다.
4. 컴퓨터 보조 접근(computer-assisted approach), 클러스터 알고리즘으로 프레임을 추출하기 때문에 과정이 객관적이지만 알고리즘이 언어를 완벽하게 이해하기 어렵다. 컴퓨터 보조 접근은 연구자가 프레임을 발견하는 것이 아니라 알고리즘이 이를 결론짓는다.

이들은 하나의 개념적인 방법론으로서, 본 연구에서는 특정 사안에 대해, 대량의 뉴스콘텐츠를 투입하여, 복합적인 프레임을 분석하고, 추출할 수 있는 구체적인 절차(process)와 그 근거의 타당성을 이론적으로 검토 해 볼 필요가 있다. 이준웅(2009)은 그의 연구에서, 일괄적인 뉴스콘텐츠의 구성과 텍스트의 나열만으로는 뉴스프레임의 추출결과가 불완전 할 수 있음을 지적하였다. 동 연구에서는 뉴스콘텐츠의 구성과 해석의 상이(相異)함에 대한 조화의 어렴움을 ‘두 개의 뿔’이라고 표현하였다. 해당 연구에서는 이에 대한 해답으로 ‘이야기(narrative)’에 대한 의미해석을 요구하였다. 구체적으로는 텍스트의 ①표현 요소 ②이야기 구조 ③함축된 가치를 포괄적으로 분석할 수 있어야 하며, 특히 ‘이야기 구조’를 골격으로 사용할 것을 제안한다. 이는 이희영, 김정기(2016)의 질적 메타분석연구와 동일한 결과이다. 이준웅 등(2018)의 연구에서는 구조모델(Topic Model)을 세가지 요소를 전제하여 구축하고(①미세먼지의 ‘위험성’을 강조하는 방식에 대한 내용, ②미세먼지의 ‘원인’을 규정하는 내용, ③미세먼지‘대책’에 대한 내용을 주제적 특성으로 갖게 될 것이라 전제하고 연구를 진행), 지침 기반의 전일적 접근법을 통하여, 기존의 ‘재난 프레임’의 복수의 프레임요소들을 동일하게 찾아내었다. 해당연구에서는 베이즈확률추정을 통하여 인위적으로 주제구조를 도출하여, 기계적으로 코딩된 연속확률분포(continuous probability distribution)알고리즘과의 대조를 통해 정확성을 검증하였다. 임종섭 등(2020)또한 동류의 알고리즘을 이용한 프레임분석을 시도하여, 유의미한 결과를 도출하였다. 해당 연구에서는 ‘해석적 묶음(interpretive package)’에 대해, 글 뭉치(corpus)에서 핵심 단어들을 알고리즘 기반의 분석 코드로 추출해 이들 간에 연관성을 의미 연결망(semantic network)과 군집(clustering)으로 추적해 프레임을 발견하고자 하였다. 두 연구는 공통적으로 주제모델(Topic Model)기반의 의미분석(Semantic Analysis)방법을 사용하였으나, 본 연구에서는 주제모델(Topic Model)을 사용하지 않는다. 이야기(narrative)구조 분석방식과 지침 기반의 전일적 접근(manual holistic approach)의 일부만을 사용하며, 워드임베딩(Word Embedding)을 사용하여 알고리즘의 추정된 프레임들을 추출할 것이다. 따라서, 본 연구에서도 비교가능한 지침구조, ‘이야기’가 필요하며, 중앙선건위원회의 선거여론조사자료<그림.3>를 바탕으로 보았을 때, 추출되는 주제유목이 아래와 같은 구조를 가지게 되리라고 가정한다.

1. 내용: 코로나및 경제위기등에 대한 위험성을 강조
2. 내용: ‘원인’으로 대통령 및 정권여당의 실정(失政)을 강조, 또는 야당의 국정발목잡기를 강조.
3. 내용: ‘대책’ 및 ‘결과’로써, 국정안정을 위해 여당지지 또는 야당지지



<그림.3>

### **워드임베딩(Word Embedding)기법과 감성분석(Sentiment Analysis)**

언어의 분석은 언어를 ‘수치’로 치환하는 것으로 시작된다. 주제모델(Topic Model)은 주제가 나타낼 특성을 빈도수를 통하여, 확률적으로 도출하는 것이다. 이 모델은 연구에서 뉴스콘텐츠의 프레임분석과 유사도 분석 및 감성분석에 시도되는 워드임베딩(Word Embedding)은 기존의 연구의 주제모델(Topic Model)이 단어의 의미를 고려하지 못한다는 단점을 보완한다. 주제모델(Topic Model)기반의 알고리즘들의 가장 큰 단점은 단어 행렬(Document-Term Matrix)에서 TF-IDF(Term Frequency-Inverse Document Frequency)를 사용한다는 것이다. 기본적으로는 베이즈추정(Bayesian inference)치를 사용하며, 이에 대한 연속확률분포(probability distribution)값을 계산한다. 따라서, 단어의 수, 글 뭉치(corpus) 크기의 증가분에 비례하여, 이를 추정하기 위한 계산복잡도도 증가한다. , ‘차원의 저주(Curse of Dimensionality)’문제가 발생한다(Bengio, Y., Ducharme, R., Vincent, P., & Jauvin, C. ,2003). 따라서, 특정 단어에 대한 인위적 조작, 가중치조작이 요구된다. 워드임베딩의 단어간의 선형적 규칙성(linear regularities among words)이, 잠재적 의미분석(Latent Semantic Analysis) 또는 잠재적 디리클레 분포(Latent Dirichlet Allocation)알고리즘보다 우위에 있다는 것은 주지의 사실이다(Mikolov, T., Chen, K., Corrado, G., & Dean, J. 2013). 워드임베딩은 단어의 특징과 유사도를 벡터(vector)기반으로 예측할 수 있기 때문에 유사 맥락의 단어를 이해할 수 있다. 이해해야 할 문장 또는 단어들은 각각의 속성을 가지고 있다. 품사, 의미, 길이등, 이러한 텍스트들의 속성을 언어정보(linguistic information)라고 하며, 이를 추출해서 차원으로 저장할 때 이를(feature representation)이라 한다. 각 단어의 표현을 동등한 차원으로 저장하는 것이 밀집표현(distributed/Dense representations)이며, 이러한 밀집표현을 이용하여 단어간의 유사성을 선형적 규칙성으로 계산, 추론 하는 것이, 본 연구에서 활용할 Word2vec이다. 단어간의 속성에는 선형적 규칙성이 존재한다. 동일한 문맥의 단어, 근접한 위치에 나오는 단어는 비슷한 의미를 가진다(distributed representations of words) (Bengio, Y., Ducharme, R., Vincent, P., & Jauvin, C. 2003). 어떤 글의 근접위치에 존재하는 단어는 단어 간의 유사도를 높게 측정할 것이고, 이에 대한 규칙성을 지속적으로 추정하고, 담보할 수 있다면, 보다 향상된 수준의 문장의 의미를 추론하여 프레임을 추론 할 수 있을 것이다(Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. ,2013; Mikolov, T., Chen, K., Corrado, G., & Dean, J. 2013).

감성분석(Sentiment Analysis)이란, 사용자가 작성한 문장의 성향을 나타내는 패턴을 이용해 텍스트 내에서 주관적인 정보를 검토하고 처리하는 기법으로 기본 작업은 문장의 극성을 긍정, 부정, 중립 등으로 나누는 것이다(홍진주·김세한·박제원·최재현,2016). 기계적으로 감성분석을 시행하는 경우는 크케 데이터마이닝과 기계학습을 사용하는 것을 생각 할 수 있다. 본 연구에서는 기계학습의 워드임베딩을 사용하여, 이용자의 ‘여당’에 긍정적인 어휘들을 사용하였는가 또는 부정정인 어휘들을 사용하였는가를 기준으로 뉴스 콘텐츠의 정치이념적 편향성을 분류한다. 각 어휘들의 워드임베딩을 통해 학습되며, ‘초평면(hyperplane)’으로 분류하기 위해 SVM(Support Vector Machine)을 사용한다. 이동엽, 조재춘, 임희석(2017)은 유사한 방법으로 소비자의 리뷰를 통해 상품의 감성을 분석하였고, 홍진주 등(2016)은 악성댓글 탐지기법을 착안하였다. 따라서, 본 연구에서는 워드임베딩과 SVM을 활용하여, 뉴스콘텐츠 정보의 프레임을 추출하고, 정치적 이념의 편향성을 각 두 사안, ‘코로나19사태’뉴스는 ‘제21대 대한민국 총선’를 언론사별로 구분 할 것이다.

## **연구방법**

### **연구대상**

#### **연구대상**

빅카인즈로(<https://www.bigkinds.or.kr/>)로 부터 추출한 뉴스콘텐츠를 연구의 대상으로 한다.

구체적인 대상의 조건은 아래와 같다.

1. 공통조건
2. 기간: 2019년12월1일∼2020년6월30일
3. 형식: 모든 기사형식
4. 대상언론: 전국단위 일간지(경제지 및 전문지포함), 지상파방송사

(조선일보, 경향신문, 국민일보, 내일신문, 동아일보, 문화일보, 한국일보, 중앙일보, 한겨례신문, 서울일보, 매일경제, 서울경제, 머니투데이, 아시아경제, 아주경제, 한국경제, 헤럴드경제, 파이낸셜뉴스, 디지털타임스, 전자신문, KBS, SBS, MBC)

1. ‘제21대 대한민국 총선’ 뉴스콘텐츠 조건: Query = (총선 OR 국회의원선거 OR 21대 선거 OR 국회의원 선거)
2. ‘코로나19사태’ 뉴스콘텐츠 조건: Query = (코로나 OR 우한 OR COVID OR COVID-19)

#### **취득 데이터 구조(Low Data Structure)**

뉴스 식별자, 일자, 언론사, 기고자, 제목, 통합 분류1, 통합 분류2, 통합 분류3, 사건/사고 분류1, 사건/사고 분류2, 사건/사고 분류3, 인물, 위치, 기관, 키워드, 특성추출, 본문, URL, 분석제외 여부

### 연구도구

뉴스콘텐츠의 취득 및 처리, 결과도출은 모두 컴퓨터 프로그램의 알고리즘으로 진행된다.

Jupyter Notebook 기반으로 Program언어로는 Python을 이용하여, Word2vec과 SVM을 구현.

Word2vec의 예측기반(Prediction Based Embedding)알고리즘을 사용하므로, 단어사용의 빈도수를 사용하지 않으며, 특정한 단어의 좌우의 50단어까지 유의단어로 차원을 할당하여, 다른 단어들과 차원을 중복시켜, 유의성에 대한 선형성을 도출한다. SVM의 경우, 주어진 변수의 차원을 2차원평면으로 치환하여 최적화된 초평면으로 귀속시킨다.

#### **Word2vec**

1. CBOW (Continuous Bag of Words)
2. Vector size=100 dimension
3. Window=R10, L10, total 20 token
4. frequency=10 times loop
5. CPU=4Q
6. Training count=100 times loop

#### **SVM (Support Vector Machine)**

1. 초평면 2차원분석
2. 로지스틱 회귀분석

### **연구절차 및 자료수집**

1. 분석대상 뉴스콘텐츠는 빅카인즈로(<https://www.bigkinds.or.kr/>)부터 추출한다.
2. 추출 대상뉴스콘텐츠는 ‘제21대 대한민국 총선’에 관련된 키워드 조합이다.
3. 추출 대상뉴스콘텐츠는 ‘코로나19사태’에 관련된 키워드 조합이다.
4. 전처리 과정을 통하여, 2)의 데이터를 정규화한다.
5. 전처리 과정을 통하여, 3)의 데이터를 정규화한다.
6. 4)의 데이터는 학습데이터(Training Data Set) 모델이다.
7. 5)의 데이터는 검증데이터(Validation Data Set) 모델이다.
8. 4)의 테이터는 Word2vec을 통하여 학습되며, 밀집표현(distributed/Dense representations)으로 생성된다.
9. 5)의 테이터는 Word2vec을 통하여 학습되며, 밀집표현(distributed/Dense representations)으로 생성된다.
10. 8)과 9)의 데이터에서 상위어휘들을 추출한다.
11. 10)의 데이터는 각 사안의 프레임이다.
12. 8)과 9)의 데이터에서 SVM을 통한 유사도를 측정한다.
13. 8)의 데이터에서 이야기구조모델 통한 유사도를 측정한다.
14. 9)의 데이터에서 이야기구조모델 통한 유사도를 측정한다.
15. 12), 13), 14)의 유사도가 비슷하다면, 프레임은 정치이념적 이라고 할 수 있다.
16. 13)의 유사도가 상위값을 가질 경우 ‘제21대 대한민국 총선’뉴스는 진보성향이다.
17. 14)의 유사도가 상위값을 가질 경우 ‘코로나19사태’뉴스는 진보성향이다.

### **자료분석방법**

#### **각 언론이 생산하는 ‘제21대 대한민국 총선’뉴스콘텐츠의 프레임 추출**

‘위험’, ‘원인’, ‘대책’에 관련된 유의어를 추출후, 밀집도를 표현하였을 때의 밀도함수 표현을 수치 및 그래프로 표현. 상위100위의 밀도를 가진 ‘단어’의 조합을 평가하여, 프레임을 추출.

#### **각 언론이 생산하는 ‘코로나19사태’뉴스콘텐츠의 프레임 추출**

‘위험’, ‘원인’, ‘대책’에 관련된 유의어를 추출후, 밀집도를 표현하였을 때의 밀도함수 표현을 수치 및 그래프로 표현. 상위100위의 밀도를 가진 ‘단어’의 조합을 평가하여, 프레임을 추출.

#### **각 언론이 생산하는 뉴스콘텐츠에서 추출한 프레임의 유사도 분석**

‘제21대 대한민국 총선’ 와 ‘코로나19사태’에서 추출된 상위100의 유사도를 SVM 으로 도출하며, 보조적으로 로지스틱 회귀분석을 이용하여 검증한다.

#### **각 언론이 생산하는 뉴스콘텐츠에서 추출한 프레임의 ‘국정안정’을 지지도 분석**

‘제21대 대한민국 총선’ 와 ‘코로나19사태’에서 추출된 상위100의 프레임이 ‘국정지지’에대한 긍정적단어로 구성되는 지를 분석한다.

### **연구설계**

Dataset

Dataset

Total News Contents

(and 21st South Korea General Election)

Total News Contents

(COVID-19)

(n)Dimension Word Embedding Space

(n)Dimension Word Embedding Space

News Contents of Each News Media (and 21st South Korea General Election)

News Contents of Each News Media (COVID-19)

Lean

Lean

Vectorized Word

Vectorized Word

Extract Similarity

Extract Similarity

Representation

Extract Similarity

Sentiment Analysis

## **연구결과 – 예상되는 연구결과**

1. ‘제21대 대한민국 총선’뉴스는 정치이념적 프레임이 추출될 것이다.
2. ‘코로나19사태’뉴스에서도1차 또는 2차적 으로 ‘제21대 대한민국 총선’뉴스 프레임과 동일한 프레임이 추출될 것이다.
3. 전체적인 뉴스프레임은 ‘국정안정’일 것이다.

# [ 부 록 ]

## **참고문헌**

엄현식·강재원(2019). 뉴스기사의 프레임이 이용자의 기사 신뢰도 인식에 미치는 영향. 사회과학연구, 26 1, 187-205.

현기득·서미혜. (2019). 한국 정파언론 환경의 특수성은 보수와 진보 수용자의 매체 태도와 이용에 차별적 영향을 미치는가?. 한국언론학보, 63(2), 46-76.

최영재(2011). 분열 정치와 분열 언론 그리고 분열 여론의 악순환 이론. 한국언론학회 심포지움 및 세미나, 137-152.

민영(2016). 신뢰의 조건 : 저널리즘 전문성과 정파적 편향성이 언론 신뢰와 정치 신뢰에 미치는 영향. 한국언론학보, 60-6, 127-156.

김세은·심훈·박영흠(2018). 보도의 신뢰 하락 원인 분석과 책임 저널리즘을 통한 방송 보도의 신뢰성 확보 방안 연구. 방통융합정책연구, KCC-2018-39.

윤성옥(2018). 가짜뉴스의 개념과 범위에 관한 논의. 언론과법, 17(1), 51-84.

이미나·구교태(2019). 미디어 이용형태와 정치심리변인들이 유권자 투표에 미치는 영향 : 6 · 13 지방선거를 중심으로. 언론과학연구, 19(4), 158-189.

염정윤·정세훈(2019). 가짜뉴스 노출과 전파에 영향을 미치는 요인. 한국언론학보, 63(1), 7-45.

김창숙(2019). 사실 확인인가, 사실 만들기인가. 이화여자대학교 대학원 박사학위논문.

임종섭·김진희(2020). 알고리즘 기반의 프레임 추출로 본 정부와 언론의 원자력 프레임 경쟁 양상 : 프레임 자동 추출 기법의 효율성과 타당성 연구. 언론과학연구, 20-1, 151-190.

이희영·김정기(2016). 질적 메타분석을 통한 뉴스프레임의 유형. 한국언론학보, 60-4, 7-38.

김병곤(2011). 한국 보수주의의 이념적 특징. 역사비평, 8-35.

이준웅(2009). 뉴스 틀 짓기 연구의 두 개의 뿔. 커뮤니케이션 이론, 5-1, 123-166.

이준웅·김성희(2018). 미세먼지 재해 보도의 프레임 분석. 한국언론학보, 62(4), 125-158.

이준웅(2017). 가짜 뉴스와 사실확인 보도. 관훈저널, 143, 33-40.

강범일·송민·조화순(2013). 토픽 모델링을 이용한 신문 자료의 오피니언 마이닝에 대한 연구. 한국문헌정보학회지, 47(4), 315-334.

이동엽·조재춘·임희석(2017). 워드 임베딩을 이용한 아마존 패션 상품 리뷰의 사용자 감성 분석. ,8(4), 1-8.

홍진주·김세한·박제원·최재현(2016). 감성분석과 SVM을 이용한 인터넷 악성댓글 탐지 기법. 한국정보통신학회논문지, 20-2, 260-267.

김하나(2017). 日本言論の編輯理念 [不偏不党] の 歴史的変化に関する研究-権力と言論の関係を中心に (Doctoral dissertation, 서울대학교 대학원).

鶴田達成(2018). 放送事業規制の日米比較と日本の問題点. 紀要, 61, 1-10.

稲積重幸(2011). 表現の自由の価値に関する一再考: 聞き手と話し手と政府言論. 札幌大学総合論叢, 31, 9-29.

Bengio, Y., Ducharme, R., Vincent, P., & Jauvin, C. (2003). A neural probabilistic language model. *Journal of machine learning research*, *3*(Feb), 1137-1155.

Robert M. Entman, Framing: Toward Clarification of a Fractured Paradigm, *Journal of Communication*, Volume 43, Issue 4, December 1993, Pages 51–58

Carragee, K. M., & Roefs, W. (2004). The neglect of power in recent framing research. *Journal of communication*, *54*(2), 214-233.

White, Jensen Kael, "A neural network for uncertainty anticipation and information seeking" (2019). Arts & Sciences Electronic Theses and Dissertations. 1764.

Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts. 2011. Learning word vectors for sentiment analysis. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1 (HLT ’11). Association for Computational Linguistics, USA, 142–150.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems* (pp. 3111-3119).

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Aker, Ahmet & Kevin, Vincentius & Bontcheva, Kalina. (2019). Predicting News Source Credibility. 10.36370/tto.2019.5.

루트비히 비트겐슈타인(1922), 논리 – 철학논고. 책세상, 이영철역, 2010

남시욱(1997). 체험적 기자론. 나남, 1997

Karl Jaspers. (1935). REASON AND EXISTENZ, FIVE LECTURES, Translated with an introduction by William Earle, Copyright 1955 by The Noonday Press, First Paper Edition (1957)

Karl Marx. (1867). Capital A Critique of Political Economy Volume I Book One: The Process of Production of Capital First published: in German in 1867, English edition first published in 1887; Source: First English edition of 1887 (4th German edition changes included as indicated) with some modernization of spelling; Publisher: Progress Publishers, Moscow, USSR; Translated: Samuel Moore and Edward Aveling, edited by Frederick Engels; Transcribed: Zodiac, Hinrich Kuhls, Allan Thurrott, Bill McDorman, Bert Schultz and Martha Gimenez (1995-1996); Proofed: by Andy Blunden and Chris Clayton (2008), Mark Harris (2010), Dave Allinson (2015).

Reuters Institute Digital News Report 2019, Reuters Institute for the Study of Journalism(<http://www.digitalnewsreport.org/survey/2019/south-korea-2019/>)

Reuters Institute Digital News Report 2018, Reuters Institute for the Study of Journalism(<http://www.digitalnewsreport.org/survey/2018/south-korea-2018/>)

Reuters Institute Digital News Report 2017, Reuters Institute for the Study of Reuters Institute Digital News Report 2019. Journalism(<http://www.digitalnewsreport.org/survey/2017/south-korea-2017/>)

Reuters Institute Digital News Report 2016, Reuters Institute for the Study of Journalism(<http://www.digitalnewsreport.org/survey/2016/south-korea-2016/>)

Navigating the ‘infodemic’: how people in six countries access and rate news and information about coronavirus, Reuters Institute for the Study of Reuters Institute(<https://reutersinstitute.politics.ox.ac.uk/infodemic-how-people-six-countries-access-and-rate-news-and-information-about-coronavirus>)

2020 WORLD PRESS FREEDOM INDEX, 국경없는기자회(RSF) (<https://rsf.org/en/ranking>)