

동적 토픽 모델링과 감성 분석을 이용한 COVID-19 구간별 비대면 근무 부정요인 검출에 관한 연구

이선민* · 천세진** · 박상언*** · 이태욱**** · 김우주*****

〈목 차〉

I. 서론	IV. 연구결과
II. 선행연구	4.1 데이터 분석
2.1 설문조사 기반의 부정요인 검출 연구	4.2 단어 빈도분석
2.2 소셜 미디어 분석 관련 연구	4.3 동적 토픽 모델링을 이용한 추이분석
III. 연구방법	4.4 트윗에 대한 감성 분석
3.1 데이터 수집	4.5 회귀분석
3.2 데이터 전처리	4.6 시사점
3.3 동적 토픽분석	V. 결 론
3.4 감성 분석	참고문헌
3.5 회귀분석	<Abstract>

I. 서론

코로나 바이러스(COVID-19, Corona Virus Disease-19)¹⁾의 대유행은 국내외 수많은 기업의 근무환경을 비대면으로 전환하는데 큰 영향을 끼쳤다. 사무실이 아닌 자택 또는 별도의 업

무공간에서 진행되는 비대면 근무는 감염 전파력이 강한 COVID-19 상황에 대응하기 위해 지속적으로 확산되고 있다. 그러나 비대면 근무를 정착했다고 하기에는 아직 이르다. 비대면 근무를 경험해본 근로자 1,000명 대상으로 실시한 미래전망 설문조사 결과(한국기업데이터, 2020)²⁾

* 연세대학교 산업공학과, imsoph4@yonsei.ac.kr(주저자)
** 동아대학교 컴퓨터공학과, sjchun@dau.ac.kr
*** 경기대학교 경영정보전공, supark@kgu.ac.kr
**** 연세대학교 산업공학과, dlxodnr22@yonsei.ac.kr
***** 연세대학교 산업공학과, wkim@yonsei.ac.kr(교신저자)

1) WHO(World Health Organization), Health Topics, COVID-19, Corona Virus Disease, 2020
<https://www.who.int/health-topics/coronavirus>

2) “언택트(Untact) 시대 미래전망 설문조사”, 한국기업데이터 R&C센터, 2020

에 의하면, 만족도가 현저히 낮은 것으로 나타났다. 기업이 비대면 근무방식에 대한 대응 전략을 체계적으로 마련하고 있는가라는 질문에 ‘그렇다’고 답한 비율은 30%에 그쳐 비대면 근무환경은 아직 미흡한 상황으로 해석된다. 오라클 자문회사인 워크플레이스 인텔리전스(Workplace Intelligence, 2020)³⁾의 업무환경과 AI 보고서에 따르면, 국내외 수많은 근로자는 COVID-19 대유행으로 인해 전례가 없는 높은 스트레스를 겪고 있는 것으로 나타났다. 사무실 출퇴근이 주는 COVID-19 감염에 대한 우려, 불안, 우울증 등이 부정적인 심리적 상태에 영향을 끼치고 있다. 그럼에도 불구하고, 국내 근로자의 절반 이상은 대면 근무를 선호하고 있다. 국외 근로자가 대면 근무를 선호하는 비율은 평균 38%임에 반해, 국내 근로자는 60%에 수준이다. 주요 원인으로서는 가정과 직장 생활의 구분 모호성에서 발생하는 어려움으로, 수면 부족, 신체건강 악화, 가정에서의 행복 감소, 가족 관계 문제, 친구와의 고립 등이 있다. 따라서, 비대면 근무의 정착 및 활성화를 위해서는 부정요인을 명확하게 파악하여 개선할 필요가 있다.

비대면 근무 관련된 기존 연구는 일반인 또는 전문가 설문조사 방식에 의존하는 경우가 대부분이다. 비대면 서비스의 채택근무 적합요인 분석(서창적, 김효정, 2013), 스마트워크가 효능감과 조직성과에 미치는 구조관계(이진영, 2020), 채택근무 및 테크노스트레스와 생산성의 연관성 분석(정용주, 김진수, 2020)의 연구들은 모두 설문조사로 진행되었다. 이와 같은 연구는 자기 보고방식의 설문조사가 갖는 표집

편향, 설문조사 과정에서 응답자가 솔직하게 태도나 감정을 드러내지 않는 사회적 바람직성 편향, 질문내용과 무관하게 일관적으로 대답하려는 목중 경향성(Schmitt, 1994) 등의 문제점을 가지고 있다. 따라서 설문조사에 기반한 연구는 일반인들의 솔직한 감정을 파악하기에는 한계가 있다.

반면, 디지털 기술 발달 및 스마트 기기 보급의 확대로 소셜 미디어에 대한 접근이 향상되고, 실시간으로 대량 데이터의 생성 및 전파가 가능해지면서, 소셜 미디어 데이터를 활용한 연구가 활발하게 진행되고 있다. 소셜 미디어를 활용한 예로, 항공사 고객만족도 연구(이주양, 2020), 여행 트렌드 분석(최홍열, 박은영, 2019), 편의점 고객 리뷰로 긍부정 의견분석(이혜진, 최자영, 2019) 등이 있다.

본 연구에서는 소셜 미디어 데이터에 대한 토픽 모델링 및 감성 분석을 통해 COVID-19 구간에 따른 비대면 근무 영향을 파악하고, 부정요인을 검출한다. 트위터에 대상으로 비대면 근무에 관련된 일반인들의 데이터 및 감성적 정보를 수집하고, 토픽 모델링을 통해 COVID-19 구간별 데이터의 추이를 분석하고자 한다. COVID-19 구간은 발생시점인 2019년 12월 31일, 팬데믹을 선언한 2020년 3월 11일을 기준으로 COVID-19 이전, 초기, 심각기로 구간을 구분하여 변화추이를 분석한다.

연구절차는 다음과 같은 단계로 진행한다. 첫째, 트위터에 게시된 트윗 중에서 비대면 근무 관련 데이터를 수집하고, 전처리하는 과정을 거친다. 둘째, 토픽 모델링인 LDA(Latent Dirichlet Allocation)를 활용하여 토픽을 추출

3) “업무환경과 AI(AI at Work) 보고서”, 오라클 및 Workplace Intelligence LLC, 2020

하고, DTM(Dynamic Topic Modeling)으로 토픽 변화를 분석한다. 셋째, 트윗의 감성 분석을 통해 긍정과 부정을 구분하고, 회귀분석을 통해 비대면 근무에 대한 부정요인을 검출한다. 기간은 2019년 7월부터 2020년 12월까지 18개월간이며, 영문 트윗 데이터를 대상으로 한다.

연구의 주요 목적은 첫째, 소셜 미디어를 통한 비대면 근무에 대한 부정요인을 검출하기 위한 프레임워크를 제시한다. 둘째, COVID-19 발생 구간별 비대면 근무의 주요 토픽 및 단어의 변화를 제시한다. 마지막으로, COVID-19 발생이후 비대면 근무에 관련된 부정요인을 검출하여 개선을 위한 시사점을 도출하고자 한다. 팬데믹 선언이후 사회적 거리두기, 이동제한이 시행됨에 따라 장기간에 걸쳐 일어날 비즈니스 모델의 변화가 단기간에 발생하고 있다. 기업은 체계적인 준비없이 비대면 근무를 도입함에 따라 이에 따른 부작용을 분석하고, 완화할 필요가 있다. 또한, COVID-19가 끝나는 시점인, 포스트 코로나에도 비대면 근무는 일정 부분 지속될 것으로 예상되므로, 기업은 효율적인 업무 환경 구성을 위해 근로자의 의견을 지속적으로 감지하여 반영하여야 한다.

본 연구 II장에서는 비대면 근무 관련 설문조사, 토픽 모델링, 감성 분석의 선행연구를 살펴보고, III장에서는 비대면 근무 관련된 데이터 수집, 구간별 토픽 및 부정요인을 검출하기 위한 프레임워크를 제안한다. IV장에서 제안된 프레임워크를 활용한 토픽 및 감성 분석을 통해 소셜 미디어 사용자의 의견 및 변화추이를 분석하고, V장에서는 연구의 결론을 맺는다.

II. 선행연구

2.1 설문조사 기반의 부정요인 검출 연구

비대면 근무에 대한 기존 연구는 일반인 및 전문가를 대상으로 설문조사 혹은 델파이 기법에 의존한다. 이진영(2020)은 스마트워크가 자기 효능감과 업무 창의성 및 조직성과에 미치는 구조관계 연구를 위해 전문가 대상 설문조사로 분석하였다. 정용주, 김진수(2020)은 재택근무 제도 및 테크노스트레스와 생산성의 연관성을 재택근무 경험이 있는 회사원을 대상으로 설문 진행 후 가설검정을 위해 SPSS(Statistical Package for the Social Sciences), AMOS(구조방정식 모델링) 프로그램을 통해 신뢰도, 요인, 구조방정식으로 효과분석을 진행하였다. 박일경(2019)는 재택형 유연근무제의 실태에 관한 연구를 위해 업무분야(영업판매, 예술, 교육, 제조업)별 재택근무자 대상 설문진행 후 SPSS를 활용하여 통계적 분석을 진행하였다. Teresa Galanti(2021)는 COVID-19 팬데믹 상황이 재택근무에 주는 사회적 고립, 근무 자율성 및 리더십 연구를 위해 재택근로자를 대상으로 온라인 설문을 진행하였다. Bondarenko Olena(2021)은 기업인 대상 설문을 기반으로 COVID-19 팬데믹이 주는 재택근무 활성화 및 가속화를 통한 기업의 기술적 개선을 연구했다. 서창적, 김효정(2013)은 델파이 기법을 이용하여 비대면 서비스 재택근무의 적합요인을 분석하였다.

설문조사 기반 연구는 자기 보고방식의 설문조사가 갖는 표집편향, 설문조사 과정에서 응답자가 솔직하게 태도나 감정을 드러내지 않는

사회적 바람직성, 질문내용과 무관하게 일관적으로 대답하려는 목중 경향성(Schmitt, 1994)이 문제가 되고 있다. 설문조사에 기반한 연구만으로는 일반인들의 실질적인 이슈를 파악하여 개선하기에는 한계가 있다. 설문조사는 일반인 또는 전문가를 대상으로 설문조사서를 작성하여 조사하고 결과를 취합하여 분석하는 과정을 거치며, 전체 구성원에 대해 조사가 현실적으로 어려우므로 일부 표본을 선정하여 가설을 검증하게 된다. 설문조사를 진행한 이후, 대상 또는 질문을 확대하거나, 주기적으로 조사하는 경우, 설문조사를 재수행하는데 장기간의 시간과 비용이 발생하게 된다.

반면, 소셜 미디어를 통한 조사는 일반인이 온라인상 게시한 내용을 대상으로 분석하므로 조사기간 및 비용이 감소되며, 대상 및 질문을 확대하거나 주기적인 분석이 가능하다. 따라서 설문조사 취합결과를 소셜 미디어 내용으로 대체하여 분석할 경우, 실시간으로 데이터 추출이 가능하며 대상자 및 조사내용을 다양하게 분석 가능하다는 장점이 있다.

2.2 소셜 미디어 분석 관련 연구

디지털 기술 발달 및 스마트 기기 확대로 소셜 미디어(SNS, Social Network Service)에 대한 접근이 확대되고, 실시간으로 대량의 데이터 생성 및 전파가 가능한 소셜 미디어를 활용한 연구가 활발하게 진행되고 있다. 그 예로, 이주양(2020)은 소셜 미디어 감성 분석을 활용하여 항공사 고객 만족도에 대한 연구를 진행하여 대형 항공사의 한국어, 영어, 일본어, 중국어로 작성된 트위터 및 웨이보 SNS를 분석하여 언

어권별 만족수준을 분석하였다. 최홍열, 박은영(2019)은 나홀로 여행 트렌드에 대해 네이버와 다음의 블로그, 카페를 대상으로 텍스트 마이닝 기법으로 키워드를 도출하여 전반적인 인식 및 트렌드를 분석하였다. 이혜진, 최자영(2019)은 편의점을 이용하는 고객 트위터의 리뷰를 활용하여 감성 분석을 통해 긍부정 의견을 분석하였다. Mostafa(2013)는 소셜 네트워크에 게시된 온라인 리뷰의 감정적 분석을 통해 브랜드에 대한 소비자의 감정을 분석하였다.

Archak(2011)은 온라인 리뷰 마이닝을 사용하여 가격 경쟁력을 결정하는 제품특성을 식별하였으며, Hu, Chen, Chou(2017)는 호텔 리뷰를 수집하고 댓글 작성자, 댓글시간 및 의견을 사용하여 주요 문장을 추출했다. 제품혁신 및 갱신과 관련된 아이디어를 얻기 위해 Christensen 등(2017)은 특정제품과 관련된 온라인 커뮤니티의 게시물과 제품 아이디어와 관련된 분류된 문장을 수집했다. 이와 같이 실시간 데이터를 활용하여 소셜 미디어의 일반인 의견을 분석하기 위한 연구는 다양한 분야로 확산되고 있다.

2.2.1 토픽 모델링 분석에 관한 연구

소셜 미디어 의견의 분석을 위해 텍스트 마이닝(Pang&Lee, 2008) 기법이 활발하게 활용되고 있다(O'Reilly, 2007). 텍스트 마이닝은 구조화되지 않은 대용량 문서에서 의미 있는 패턴이나 정보를 추출한다. 이 중 LDA는 토픽 모델링 기법으로, 확률 모델인 디리클레 분포추정(Blei, 2003)을 사용하여 문서의 여러 토픽, 전체 문서의 토픽 비율 및 특정 단어가 토픽에 포함될 확률을 찾는다(Teh, 2006). Tetlock(2007)

은 미디어 심리가 주식시장에 미치는 영향을 LDA 기법으로 분석하였고, Grimmer(2010)는 미국 상원의원의 연설문과 보도자료에서 토픽을 추출하여 정책의 우선순위를 식별하고 분석하는데 사용하였다. Wang 등(2020)는 지식경제 사회에서 경쟁력 유지를 위해 기업기술 경쟁력 평가모델을 LDA를 활용하여 분석하였다. 차영란(2021)은 코로나 이전과 이후 4차 산업혁명에 관한 광고의 변화를 LDA를 통해 분석하였다.

토픽 모델링인 LDA는 문서의 집합으로부터 어떤 토픽이 존재하는지를 알아내기 위한 알고리즘으로, 단어가 특정토픽에 존재할 확률과 문서에 특정 토픽이 존재할 확률을 결합확률로 추정하여 토픽을 추출한다. LDA는 시간에 따른 변화를 반영하지 않고, 문서의 토픽을 추론할 때 사용한다(고은지, 최선영, 2021). 시간에 따른 토픽의 트렌드를 파악하기 위해 구간별로 나누어 LDA를 실시하게 되면, 각 구간별 토픽이 서로 독립적이므로 시간에 따른 변화를 추적하기가 어렵다. 반면 전체 구간에 대해 LDA를 실시하면 특정 구간에서 어떤 토픽은 사라지고 다른 토픽이 출현해도 이를 파악하기가 어렵다.

DTM은 이를 보완하기 위한 방법론으로, 각 토픽의 시간에 따른 경향성을 가정하고 이에 따라 토픽을 추출함으로써 시간에 따른 토픽의 변화를 체계적으로 파악할 수 있게 한다. DTM은 토픽이 고정된 시간 간격으로 순차적 문서 처리를 가능하게 하는 LDA의 확장된 모델이다. LDA는 시간추이를 반영하지 않은 문서 집합의 토픽을 확률적으로 추론하고 있으며, 시간 변화에 따라 토픽이 달라질 가능성이 높은 데

이터의 경우는 순차적 LDA인, DTM을 통해 토픽 추이를 확률로 구할 수 있다. 강찬희(2021)는 소셜 미디어상 형성된 코로나19의 사회적 이슈를 분석하기 위해 DTM의 동적 토픽 모델링을 활용하여 시계열적으로 이슈토픽을 파악하였다. 우창우(2020)는 데이터, 네트워크, AI 분야의 연구동향 분석을 위해 국내 연구개발 지원하는 국가연구 개발사업과제 중 ICT 분야의 시계열 분석을 위해 DTM 모델을 적용하였다. 토픽 모델링은 대량의 문서 패턴을 파악하여 유사한 의미를 갖는 단어를 묶어서 토픽을 추론하기 위해 최근 데이터 분석 분야에 활발하게 사용되고 있다.

2.2.2 감성 분석에 관한 연구

감성 분석(Sentiment analysis)은 감정이 식별된 텍스트의 집합을 사용하여 모델을 구성하고 긍정적인 혹은 부정적인 감정을 생성된 모델을 통해 예측한다. Zeng, Duan, Wang 및 Wu(2011)는 다중혼합 토픽모델을 사용하여 문서에서 상세한 토픽을 추출하는 방법을 제안했고, Xianghua, Guo, Yanyan, Zhiqiang(2013)은 LDA 모델을 사용하여 토픽을 추출하고 HowNet 어휘를 사용하여 감정을 분류하는 연구를 진행하였다. Quan, Ren(2014)은 말집합을 사용하여 제품 기능에 대한 감정을 추출하는 방법을 제안하였다. Jin, Ji, Gu(2015)는 사용자 리뷰에서 제품의 특성을 나타내는 문장을 추출하고 감정적 분석을 수행하였다.

LSTM(Long Short-Term Memory)은 시계열 데이터에 대해 데이터의 이전 혹은 이후 정보에 따른 순서 영향을 고려한 심층신경망 모델로 이전 단락에서 언급된 문제들을 해결하는데

유용하다. 트위터에 대한 감성 분석으로 가장 많이 쓰이는 방법으로, Bi-LSTM(Bidirectional LSTM)은 트위터내 이전 혹은 이후 단어에 대한 정보를 인공신경망에 보존하여 단어간 관계 정보를 반영할 수 있다. Cliche(2017)는 트위터 전후 정보 맥락을 반영하여 감성 분석을 진행하기 위해 데이터셋을 통해 LSTM을 적용하였다. Shahid, Zameer, Muneeb(2020)는 COVID-19로 인한 변화를 예측하기 위해 단기 예측모델로 LSTM과 Bi-LSTM을 활용하였다. Barnes, Klinger, Walde(2017)는 벤치마크 데이터셋을 통해 LSTM과 Bi-LSTM 모델의 성능을 비교한 결과, 전반적으로 Bi-LSTM 성능이 우수하게 나타났고, 미세하게 정의된 데이터셋에 한해서는 유사한 수준을 나타냈다. 특히, Bi-LSTM은 문장의 길이가 트위터와 같이 일정하지 않은 경우에도 효과적으로 처리할 수 있으며, 과거와 미래 정보를 정확하게 파악 가능하다는 장점이 있다(Limsopatham, 2016).

감성분류 모델에서는 텍스트 단어간의 관계를 인지하여 문서 작성자의 감정을 추출하는 기술 적용으로, 특정토픽에 대해 긍정 또는 부정의견을 정량화 할 수 있다. 최근 기업에서는 브랜드 이미지에 대한 부정적 의견을 모니터링하거나, 신제품에 대한 호응을 측정하기 위해

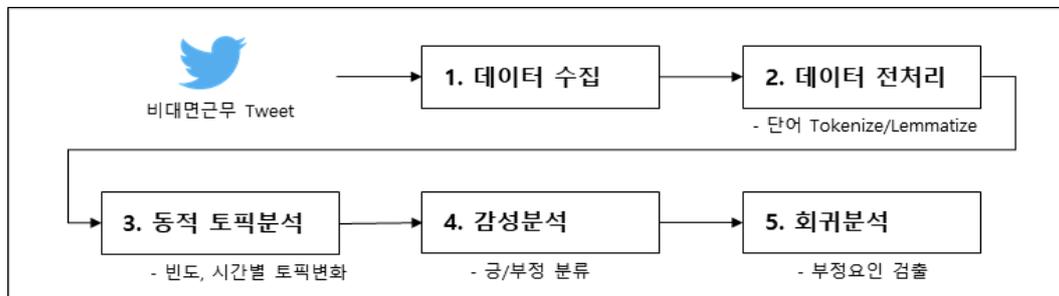
감성분석 방법론을 활용하고 있다.

Ⅲ. 연구방법

제안하는 비대면 근무 관련 토픽 기반의 감성 분석 연구단계는 <그림 1>과 같이 트위터 데이터 수집 및 정제를 통해 토픽 및 감성 분석을 진행한다. 먼저 2019년 7월부터 2020년 12월까지 18개월간의 데이터를 수집하고, 2단계는 단어를 변환, 분리 및 제거하여 전처리한 후, 3단계는 동적 토픽분석을 진행하고, 4단계는 감성분석을 통해 긍정과 부정의견을 분류하고, 5단계는 회귀분석으로 부정요인을 검출하였다.

3.1 데이터 수집

데이터를 수집하는 단계에서는 대표적인 소셜 미디어 사이트 중 하나인 트위터로부터 파이썬 기반의 웹 크롤러인 TWINT API(Twitter Intelligence Tool, <https://github.com/twint-project/twint>)를 이용하여 비대면 근무를 포함한 트윗을 수집하였다. DTM을 수행하기 위해 Tomotopy API(<https://bab2min.github.io/tomotopy>)을 이용하여 수집된 트위터의 코퍼스로부터



<그림 1> 연구모형

터 모델을 생성하였다. 두 모델 생성을 위한 파라미터는 주제의 일관성(Coherence)와 혼란도(Perplexity)를 고려하여 최적화되었다. 감성분석의 모델 생성을 위해 Tensorflow (<https://www.tensorflow.org/>)를 사용하여 Bi-LSTM 기반 모델을 학습 및 생성하였다. 비대면 근무는 재택근무, 원격근무, 통신근로, 스마트워크⁴⁾ 용어로 정의되며, 이를 정리한 문헌(김경태, 2019)을 참고하여 이 4개 단어가 포함된 영문 트윗터를 수집하였다. 트윗터 데이터는 사용자 및 트윗건수의 변화추이를 분석하기 위하여 사용자 ID, 트윗 게재시간, 내용은 수집하고, 개인 정보, 사진, Link 관련 정보는 성능향상을 위해 제외하였다. 또한 고유의 데이터만 남도록 중복된 트윗도 제외하였다.

COVID-19 이후에 기업은 비대면 근무를 본격화하기 시작하였으므로 본 연구에서는 COVID-19 관련된 시점을 3개의 구간으로 구분하여 비교하였다. COVID-19가 세계보건기구(WHO)에 최초 보고된 2019년 12월 31일, COVID-19에 대해 세계적 대유행, 즉 사람들이 면역력을 갖고 있지 않은 새로운 질병이 예상 이상으로 전 세계에 퍼지는 경우를 말하는 '팬데믹'을 선언한 2020년 3월 11일을 기점으로 하였다. 분석구간은 발생 시점을 기준으로 COVID-19 이전, 초기, 심각기로 구분하여 COVID-19 이전은 2019년 7월 1일에서 12월 31일, 초기는 2020년 1월 1일에서 3월 11일, 심각기는 2020년 3월 12일에서 12월 31일로 구분하였다.

3.2 데이터 전처리

데이터를 전처리하는 단계에서는 트윗의 특수문자, 이모티콘 및 불용어를 제거하고, 단어 집합을 구성하였다. 불용어 제거 과정에서는 @태그, #태그, URL, 광고를 포함하여 제거하였다. 단어 토큰화 과정에서는 명사만 추출한 후, 표제어 추출(Lemmatization) 과정에서 기본사 전형 단어로 변환하였다. 단어 정제과정에서는 트윗에 사용되는 약어를 변환하여 최종 정제된 데이터를 생성하였다. N-gram 데이터 구성과정에서는 N그램을 '_'으로 연결하여 Unigram, Bi-gram 데이터를 혼합한 명사를 구성하였다.

3.3 동적 토픽분석

동적 토픽분석 단계에서는 먼저 트윗 데이터 전체의 특성 및 구조적 관계를 파악하기 위해 탐색적 데이터 분석인 EDA(Exploratory Data Analysis) 방법을 활용하여 데이터 구조를 분석하였다. 일반적인 통계모델과 같이 단어 빈도분석 및 확률모델을 포함(O'Connor, 2011)하여 전반적인 현황을 파악하였다. 수집된 데이터 건수, 월별 증감, 상위빈도 단어, COVID-19 구간별 빈도 증가 및 감소한 단어를 검출하고, COVID-19 이후 신규로 출현한 단어도 분석하였다.

토픽 모델링인 잠재 디리클레 할당(LDA) 방식을 통해 비대면 근무 관련 트윗 데이터의 토픽을 검출하였다. 토픽 모델링은 자연어 처리에서 사용되는 기법으로, 하나의 문서에 다양한 주제가 있다고 가정하고, 문서 내 속한 단어로

4) 영문 : Work from home, Remote work, Telework, Smart work

각 문서가 토픽을 얼마나 포함하는지 계산을 한다. 각 트윗은 토픽의 확률적 혼합으로 구성되며, 각 토픽에서 추출된 단어로 트윗 텍스트의 구조 파악이 가능하다. 본 연구에서는 LDA를 시간에 따른 토픽 분석에 적합하게 확장한 DTM 기법을 포함하여, COVID-19 구간별 시간에 따른 토픽 변화추이를 분석하였다.

3.4 감성 분석

감성 분석 단계에서는 트윗의 긍정 및 부정에 대한 분류를 진행하였다. 감성 분석은 감정의 상태 또는 주관적 정보의 패턴을 정량화하여 데이터에 존재하는 성향을 식별하고 분석하는 기술이다. 먼저 감성 분석을 위한 분류기로 잘 알려진 딥러닝 모형인 Bi-LSTM 모형을 사용하였다. Barnes et al.(2017)에 따르면 SST-fine, SST-binary 등의 다양한 데이터 셋에 대해 CNN, LSTM, Bi-LSTM과 같은 딥러닝 모형의 성능을 테스트한 결과, Bi-LSTM이 평균적으로 가장 우수한 성능을 보였다. 또한, 트윗을 대상으로 한 감성 분석 성능 비교 연구인 Vedurumudi (2021)에서도 LSTM이 CNN에 기반한 다양한 모형과 거의 유사한 성능을 보였다. 따라서 본 연구에서는 Bi-LSTM을 감성분류 모형으로 선택하여 감성 분석을 실시하였다.

주어진 문서를 Bi-LSTM에서 사용할 워드 임베딩 벡터로 변환하기 위하여, 스탠포드대에 사전학습하여 공개 배포한 GloVe 임베딩 벡터 중 하나인 Glove.6B.50d⁵⁾를 사용하였다.

Glove.6B.50d는 60억개의 훈련된 단어, 40만 개의 고유단어로 구성되어 있으며, 임베딩 벡터는 50개 차원으로 되어 있다. 다음으로 감성 분석 분류기의 학습 데이터로는 잘 알려진 캐글 데이터셋인 Sentiment140(Go, Bhayani, Huang, 2009)을 사용하였다. Sentiment140은 트위터 API를 이용해 추출된 사용자 66만 명의 160만 건 트윗에 대해 긍정, 중립, 부정 라벨을 포함하고 있다.

3.5 회귀분석

다양한 비대면 근무 관련 각 토픽이 감성 분석의 결과로 도출된 감정 점수(Sentiment Score)에 미치는 영향 분석을 위해 COVID-19 구간별 회귀분석을 진행하였다. 회귀분석 단계에서는 종속변수인 감정 점수를 설명하는 독립변수인 토픽을 하나씩 제거하여 회귀모형을 선택할 수 있도록 단계적 다중 회귀분석(Stepwise Multi Regression)을 적용하였다. 단계적 다중 회귀분석은 독립변수를 모두 포함한 완전모형에서 설명력이 가장 작은 설명변수부터 하나씩 제거하고 각 단계에서 이미 선택된 변수의 중요도를 확인하여, 중요하지 않은 변수를 제거한다. 각 토픽의 계수에 따라 긍정요인과 부정요인으로 구분하였으며, COVID-19 구간에 따라 긍정요인과 부정요인이 어떻게 변화하는지 분석함으로써 부정요인에 대한 분석을 진행하였다.

5) Kaggle Dataset : Glove.6B.50d(<https://www.kaggle.com/watts2/glove6b50dtx>)

<표 1> 데이터 수집 현황

검색키워드	구간별 건수			
	합계	이전(6개월)	초기(3개월)	심각기(9개월)
재택근무	924,835	92,667	251,504	580,664
원격근무	519,459	59,741	96,016	363,702
통신근로	94,691	4,956	32,160	57,575
스마트워크	5,210	2,576	1,214	1,420
트윗건수 계	1,544,195	159,940	380,894	1,003,361
사용자ID 계	366,052	18,428	66,612	281,012

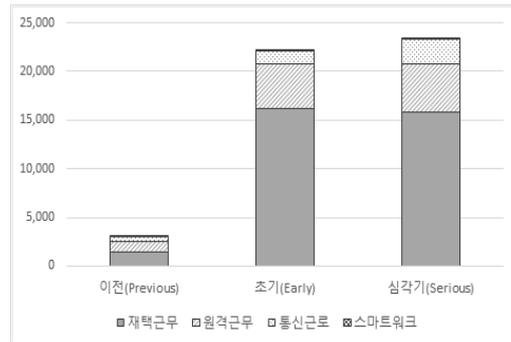
IV. 연구결과

4.1 데이터 분석

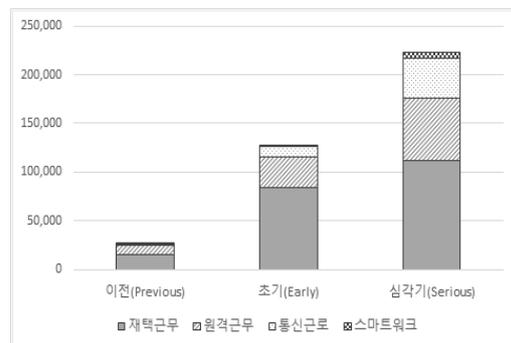
비대면 근무 관련 데이터 셋은 소셜 미디어 인 트위터 사용자 ID 36만 개, 트윗건수 154만 건을 수집하였다. 전체 트윗 건수는 <표 1>과 같이 154만 건으로, COVID-19 이전 16만 건, 초기 38만 건, 심각기 100만 건이 수집되었다. 비대면 근무유형으로는 재택근무가 92만 건으로 가장 많은 60% 비중을 차지하고 있다. COVID-19 구간별 재택근무, 원격근무, 통신근로는 모두 증가세를 보이고 있으나, 스마트워크는 오히려 감소하였다.

비대면 근무 관련된 사용자 ID의 월 평균은 <그림 2>와 같이 COVID-19 이전 대비 초기 7.2배, 심각기는 7.6배로 증가하였다. COVID-19 초기 이후 비대면 근무 관련 언급하는 일반인은 심각기에도 일정수준 유지되고 있음을 알 수 있다. 비대면 근무 관련된 트윗 건수의 월 평균은 <그림 3>과 같이 COVID-19 이전 대비 초기 4.8배, 심각기는 8.4배로 증가하였다. COVID-19 초기 단계에 비대면 근무 관련된 관심이 지속적으로 증가한 것으로 해석된다. 비대

면 근무유형별로는, COVID-19 특성상 사회적 거리두기 및 이동제한으로 인해, 재택근무가 가장 큰 증가세를 보이고 있다.



<그림 2> 월 평균 사용자 ID



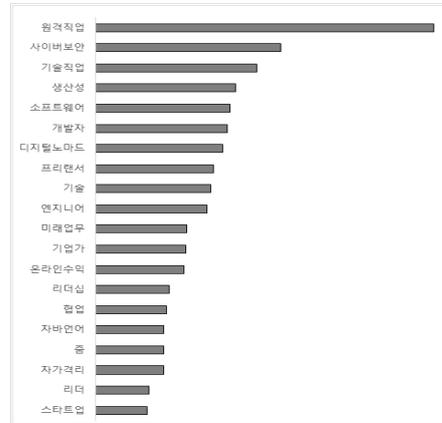
<그림 3> 월 평균 트윗건수

4.2 단어 빈도분석

4.2.1 상위빈도 단어

비대면 근무 관련 전체 트윗에 대해 상위빈도 단어, COVID-19 구간별 증가 또는 감소한 단어, 그리고 COVID-19 이후 신규 출현한 단어에 대해 분석하였다. 상위빈도 단어는 <그림 4>와 같이 원격직업, 사이버보안, 기술직업, 생산성, 소프트웨어 순으로 나타났다. 비대면 근무가 본격화됨에 따라, 원격직업, 기술직업에 대한 관심과 사이버보안, 생산성에 관련된 우려가 나타남을 알 수 있다. 또한 디지털 노마드와 같이 인터넷과 업무에 필요한 기기와 작업공간만 있으면 시간과 장소에 구애받지 않고 일하는 근로자를 나타내는 단어도 나타났다. IT에 대한 요구증가로 소프트웨어, 개발자, 엔지니어, 자바언어도 나타났다. 업무에 있어서는 리더십, 협업에 대한 필요 증가로 높은 빈도로 나타났다.

상위빈도 단어를 <표 2>와 같이 업무, 직종, IT/보안, IT 기업, 온라인사업 영역별로 분류한 결과, 업무 관련하여서는 생산성, 리더십, 협업, 비즈니스 연속성이 중요해지고 있으며, 직종으로는 디지털 기반의 업무체계 구성을 위한 개



<그림 4> 상위빈도 단어

발자, 엔지니어, 분석가, 전문가 등이 다수 언급되고 있다. IT 및 보안 관련해서는 채택근무가 활성화됨에 따라, 다양한 장소 및 기기 연결로 인한 사이버 보안의 중요성, 클라우드, 플랫폼의 필요성 및 디지털 전환에 대한 언급이 있었다. IT 기업으로는 대표적인 구글, 아마존, 애플, 마이크로소프트 그리고 협업 도구를 제공하는 슬랙이 높은 빈도로 나타났다. 그 외, 온라인을 통한 수익창출, 원격직업으로 재정적 자유에 대한 관심이 높아짐을 알 수 있다.

4.2.2 구간별 단어빈도 변화

COVID-19 구간에 따라 빈도가 증가한 단어

<표 2> 영역별 주요단어

영역	주요단어
업무	생산성, 리더십, 협업, 동료, 워크-라이프 밸런스, 비즈니스 연속성
직종	개발자, 프리랜서, 엔지니어, 기업가, 분석가, 전문가
IT/보안	사이버 보안, 기술, 자바언어, 클라우드, 플랫폼, 웨비나, 디지털 전환, 어플리케이션/앱
IT기업	구글, 아마존, 애플, 마이크로소프트, 슬랙, AWS
온라인사업	원격/기술직업, 온라인 수익, 재정적 자유

<표 3> 주요 단어변화

구분	단어
증가단어	락다운, 비즈니스 연속성, 글로벌이동성, 가상미팅, 백신, SD-WAN, IT보안, 사이버범죄, 직업검색, 제로트러스트, 가상직업, 가상문화, 업무복귀, 사이버공격, 해고, 기술직업, 비즈니스 전환, 원격업무 기회, 정보보호
감소단어	소셜 미디어로 수익창출, 체중감량 팁, 뷰티 팁, 글로벌시장, 풀타임 직업, 임금수익, 휴일, 핸디맨, 홈트레이닝, 선불, 블로그, 비즈니스 승진, 여행 팁
신규단어	격리생활, 사회적 거리두기, 자가격리, 재택 챌린지, 손씻기, 위생, 락다운 연장, 줌 미팅, 가상이벤트, 원격기술서비스, 온라인 무료캐시, 틱톡 틱즈, 구글 미트, 뉴노멀

는 <표 3>과 같다. 락다운과 해고에 대한 우려와 업무복귀에 대한 기대가 동시에 증가하고 있고, 비즈니스 연속성에 대한 고민과 이에 대한 해결책으로서의 비즈니스 전환에 대한 관심이 증가하는 것을 볼 수 있다. 또한 가상미팅이 활성화됨에 따라 IT보안, 사이버 공격, 정보보호에 대한 관심도 동시에 증가하고 있다. 빈도가 감소한 단어는 COVID-19가 심각해짐에 따라, 소셜 미디어를 통한 수익창출 기대, 풀타임 근무, 승진의 어려움이 나타났다. 재택근무 기간이 장기화됨에 따라 사회적 고립으로 인해 체중감량, 뷰티에 대한 관심은 줄고, 이동제한으로 여행의 기회도 감소된 것으로 해석된다.

COVID-19 이후 트윗에 신규로 등장한 단어는 격리생활, 사회적 거리두기, 재택 챌린지, 위생 등으로, 전염병 대유행에 관련된 단어가 나타났다. 팬데믹 선언으로 비대면 근무지원을 위한 IT 인프라 관련 해서는, 줌 미팅, 구글 미트와 같은 원격근무 도구가 출현하였고, 새로운 현상의 시작을 알리는 뉴노멀이라는 단어가 신규로 언급되었다.

4.3 동적 토픽 모델링을 이용한 추이분석

4.3.1 토픽분석

비대면 근무 관련 트윗의 토픽과, 각 토픽의 시간에 따른 추이를 분석하기 위해 LDA 기법에 기반한 동적 토픽 모델링인 DTM을 수행하였다. 먼저 전체 기간의 토픽에 대해 살펴보면, 비대면 근무 관련 토픽으로 <표 4>와 같이 비즈니스 성장, 업적, 생산성 저하에 대한 우려, 업무환경 개선을 위한 원격사무실, 가상지원 등이 추출되었다. 토픽별 주요 단어에 대해 살펴보면 먼저 비즈니스 성장과 관련하여 온디맨드 주문형, 비즈니스 전환, 아웃소싱과 같은 단어들이 COVID-19로 인한 비즈니스 형태의 변화를 나타내고 있다. 블로소득의 경우는 암호화폐, 비트코인, 블록체인과 같이 근로수당 이외에 부가적인 소득에 대한 관심의 증가를 보여준다. 이 외에 비대면 근무의 확산에 따른 변화에 대한 기회와 우려가 나타났는데, 예를 들어 “당사 글로벌 인사 책임자는 COVID-19 대응하여 신속하게 T/F를 구성하고, 원격근무가 가져올 잠재적으로 긍정 영향을 논의하고 있다6)”, “COVID-19는 미래 원격직업에 영향을 미치고 프리랜싱 경제에 큰 변화를 일으킬

것?” 등의 트윗이 이를 보여주고 있다.

비대면 근무 수행을 위해서는 IT 인프라의 지원이 수반되어야 한다. 기업에서는 비대면 근무환경을 위해 네트워크 확충, 개인노트북의 사내 인프라 접속환경 구성, 화상 커뮤니케이션 등을 지원하고 있다. 비대면 근무 관련 토픽 중에서도 IT 인프라와 관련된 토픽이 <표 5>와 같이 다수 도출되었다. AI, VR, IoT, 클라우드를 통한 디지털 트랜스포메이션이 가속화되고

있으며, 향후 업무의 변화를 나타내는 미래업무, 디지털 노마드도 언급되었다. CIO 토픽에서는 원격지에서도 근무가능한 IT 인프라 환경 구성을 위한 원격액세스, 원격지원, 어플리케이션이 언급되었다. 디지털 전환 토픽에서는 디지털 워크플레이스 구성을 위한 전문인력이 언급되었다. 하이브리드 H/W 토픽에서는 가상화, 선진화를 통한 다양한 인프라 환경 구성이 있었다.

<표 4> 비대면 근무 관련 토픽

토픽	상위단어 (주요단어)
가상지원	가상팀, 가상지원 서비스, 트레이너, 채용담당자, 링크드인, 구매자
리더십	기업가, 스타트업, 창업, 파트타임 업무, 협업 도구, 소비, 장벽
블로소득	비트코인, 암호화폐, 블록체인, 안드로이드, 디지털노마드, 부채
비즈니스 성장	On-Demand 주문형, 비즈니스 전환, 아웃소싱, 프로파일, 에이전시, 재정서비스
비즈니스 오너	시간관리, 복잡성, 팀원, 도메인, 참석, 와인, 디지털 에이전시, 채용
업적/성취	표준/기준, 분산팀, 긍정 마인드, 허브, 표현, 탈출/모면, 갑부
생산성	프로젝트 관리자, 인프라, 마켓 플레이스, 아마존, 클라우드 보안
원격사무실	랜드스케이프, 공원, 요가, 포트폴리오, 체육관, 소비, 원칙
원격직업	데이터 사이언스, SQL, 제품관리, 본사, 유튜브, 동료, 기술뉴스
유연근무	업무 트렌드, 원칙, 완충장치, 밀레니얼 세대, 행복감, 자본
자동화	기업가, 제조사, 지능, 구직자, 관리자, 의무, 기밀, 피로감
재정적 자유	유튜브, 창작, 생산성, 해킹, 하부구조
정책	자율성, 차단, 독립성, 혁신가, 분석가, 방식, 영구성
토론	글로벌, 태도, 공유, 반영, 치료, 급여, 유발
퇴직	프리랜서, 선호업무, 자기계발, 패턴, 취약성, 정치적 견해
협업	업무문화, 동기부여, 우울감, 예측, 기후, 깃허브

<표 5> IT 관련 토픽

토픽	상위단어 (주요단어)
AI/에자일 전문가	SaaS, 효율성, 민첩성, 인재채용, 전문성, 선호성, 용의성
CIO	어플리케이션, 원격액세스, 원격지원, 피드백, CISO, 악성S/W
IT 기술	개발자, 자바스크립트, DevOps, 파이썬, AI, AWS, 빅데이터

- 6) Tweet 원문 : “Our Global Chief Talent Officer discusses on our COVID-19 with task force team regarding potentially positive impacts on the remote jobs and operation in the future”, May 8, 2020
 7) Tweet 원문 : “COVID-19 will impact the remote jobs in future, Freelance Economy will get a hype”, May 26, 2020

S/W 엔지니어	CTO, CMO, 과제, 세미나, GameDev, 워크데이, 역동성
VR/IoT	스마트 작업, 다양성, 블루프린트, 데이터베이스, 라우터, 투자
기술	온라인 비즈니스, 가상사무실, 사무공간, VOIP, 경쟁, 빌링
기술직업	원격직업, 워크라이프 밸런스, S/W엔지니어, 블로거, 제출기한
디지털 노마드	중소기업, 가상직업, 유연근무, 협업업무, 업무복귀, 독립성, 단순성
디지털 전환	디지털워크플레이스, 전문가, 작업 유연성, 구글, 플러그, 사업지원
미래업무	WIFI, VDI, IT 셋업, 유연성, 대리인, 벤라이프, 아웃도어, 야외
사이버 보안	해커, 취약성, 정보보호, 구독형, 리더, 공공기관, 준비수준, 규칙
앱 플랫폼	프로듀서, 커스텀, 튜토리얼, 인스타그램, 채용기회, 보조수익
자바 개발자	원격직업, DevOps, 코딩, 마이크로소프트팀, Python, AI, AWS
클라우드 전환	마이그레이션, 파이프라인, 사내인프라, 연속성, SME
하이브리드 H/W	가상화, 선진화, 공동직업, 온라인 사업, 참조사이트, 가상지원



<그림 5> COVID-19 초기 정점 토픽

4.3.2 토픽 트렌드 분석

COVID-19 기간 동안 토픽의 추이를 보기 위해, COVID 초기와 심각기에서 정점(Peak)인 토픽들을 도출하였다. COVID-19 초기 정점인 토픽은 <그림 5>와 같이 비대면 근무 초기 IT 인프라 대비 필요, 비즈니스에 대한 우려로, CIO 및 비즈니스 성장이 나타났다. COVID-19가 장기화됨에 따라 비대면 근무 관련된 IT 인프라에 대한 지원이 안정화되고, 사용자는 근무 환경에 익숙해짐에 따라 토픽이 감소함을 알 수 있다. 또한 비대면 근무의 본격 시행으로 주

택에 대한 관심 및 최소한의 소유만 하는 미니멀리즘 토픽이 일시적으로 증가하였다. 미니멀리즘과 관련해서는 “재택에서 일하는 경우 집중력과 생산성을 유지하기는 어렵지만, Zen 근무환경을 구성하면 스트레스를 줄이고 집중할 수 있다⁸⁾”는 의견이 있어, 재택근무에 따라 자택내 근무환경에 대한 관심도 증가했음을 알 수 있다.

COVID-19 장기화에 따라 심각기 이후 정점인 토픽은, <그림 6>과 같이 재정적 자유에 대한 관심 증가, 비대면 근무지원에 대한 체계적인 정책 및 온라인 업무 시 필요한 가상지원이

8) Tweet 원문 : “It can be difficult to remain focused and productive while working from home, but by creating a zen workplace, you can reduce stress and zero in on what needs done”, Sept. 1, 2020



<그림 6> COVID-19 심각기 정점 토픽

나타났다. 재정적 자유 관련해서는, 직업, 수입에 대한 불안정성으로 인해 근무의 수당에 대한 관심이 증가하였다. 관련한 트윗의 예로, “9시 출근해서 5시 퇴근하는 노예생활에서 지쳤는가? 재정적인 자유를 찾는 길이 있다⁹⁾” 등이 있다. 정책 관련해서는, 재택근무 관련되어 근로자를 대상으로 하는 정책 및 가이드는 많지만, 비즈니스 오너를 대상으로 하는 관리방안, 근무방침은 전무하며, IT 중심의 정책이 재택근무 전반으로 확대되어야 한다는 의견이 있다. 재정적 불안정성, 가상지원에 대한 필요는 심각기가 6개월 이상 지속됨에 따라 감소함을 알 수 있다.

4.4 트윗에 대한 감성 분석

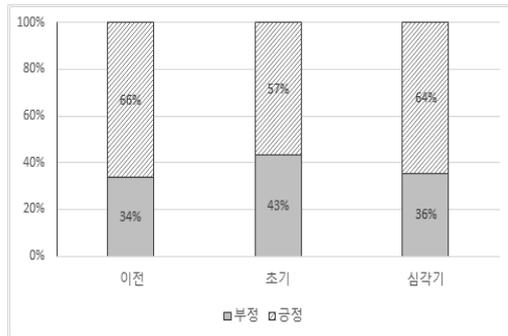
비대면 근무 관련된 트윗에서 감성분석을 한 결과, 일반인의 전체 평균은 긍정 64%, 부정 36%로 나타났다. COVID-19 구간별 부정의견은 <그림 7>과 같이 이전 34%에서 초기 43%로 증가하였다가, 심각기 36%로 감소되었다. COVID-19 초기에 비대면 근무를 전면 시행함

에 따라 재택근무 업무환경 초기 구성, 네트워크 속도 이슈, 업무와 생활을 병행함에 따른 불편함을 경험하면서 부정의견이 증가한 것으로 해석된다. “COVID-19로 인해 오늘날 수백만 명의 근로자가 익숙하지 않은 문제에 직면하게 되었으며 이는 재택근무 환경의 복잡성을 포함하고 있다. 재택근무의 어려움을 극복하고 생산성을 극대화하는 방법을 찾아야 한다”는 의견을 참조하면, COVID-19 초기에 익숙하지 않은 환경으로 인한 불편함으로 해석된다. 재택근무가 지속되면서 근로자는 원격 접속환경에 익숙해지고, 비대면 근무에 관련된 부정의견은 COVID-19 이전 수준으로 감소하였다.

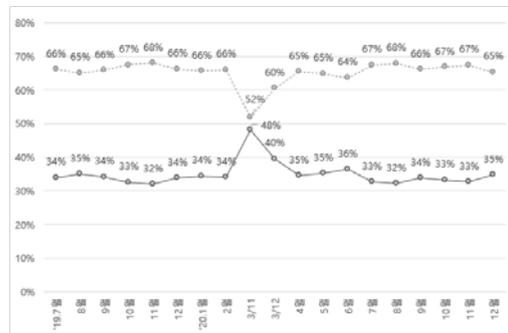
COVID-19 기간내 월별 변화를 보면, <그림 8>과 같이 2019년 7월은 34% 수준인 부정의견이 팬데믹을 선언시점인 2020년 3월 11일 일시적으로 48%로 증가하였다가 4월 이후 35%로 감소하였다. 팬데믹을 선언한 2020년 3월 11일 당일, “재택근무를 위해 집에서 노트북을 설정해 보기 전까지는 좌절의 진정한 의미를 모를 것이다¹⁰⁾”, “우리는 몇주동안 집에 있어야 한다고 들었다. 나와 같은 사람은 재택근무를

9) Tweet 원문 : “Are you tired of the 9-5 slave trade? Want to create a live you deserve? There are ways to financial freedom”, April 13, 2020

할 수 없기 때문에 무급휴가에 직면하고, 복귀할 직업도 없게 된다. 많은 사람들은 대출과 월세 지불로 인해 빚을 지게 될 것이다¹⁾”와 같은 부정적인 의견이 있다. COVID-19 팬데믹 선언으로 인한 업무와 생활방식의 변화가 일부 부정적인 영향을 미치고 있는 것으로 보인다.



<그림 7> 비대면 근무 COVID-19 구간별 감성변화



<그림 8> 비대면 근무 월별 감성변화

4.5 회귀분석

회귀분석을 위한 회귀식은 아래와 같으며, 식에서 Y는 감성 분석을 통해 도출된 문서의 감정 점수이고, X는 문서의 토픽 확률분포를 나타낸다. 즉 식은 각 토픽이 감정 점수에 미치는 영향을 표현한다고 할 수 있다.

$$Y = \alpha + \beta_1 * t_1 + \beta_2 * t_2 \dots \beta_{50} * t_{50}$$

(Y = 감정 점수, α = 회귀선 지점,
β = 회귀선 기울기, t = 토픽)

비대면 근무 관련 각 토픽이 감정 점수에 주는 영향을 회귀분석을 통해 분석한 결과, COVID-19 심각기에 부정적인 영향을 주는 토픽이 증가하였다. COVID-19가 심각해짐에 따라 비대면 근무에 대한 부정적인 영향을 주는 토픽은 <표 6>과 같이 이전 1개, 초기 9개, 심각기 14개로 점진적으로 증가하였다. COVID-19 팬데믹 이후 본격 재택근무가 시행된 심각기에는 비즈니스 성장, 업적에 대한 우려, 정책 부재에 대한 부정요인이 나타났다. 또한 커뮤니케이션의 어려움으로 토론, 원격사무실에 대해서도 부정적으로 나타났다.

COVID-19 구간에 따라 각 토픽의 감성변화를 분석해보면, <표 7>과 같이 비대면 근무의 장점인 업무 생산성, 유연성에 대한 토픽은 전 구간에서 긍정적인 영향을 주고 있다. 예를 들어 “원격근무는 높은 업무 유연성과 독립성을

10) Tweet 원문 : “You don't know the true meaning of frustration until you've tried to set your laptop up to work from home”, Mar. 11, 2020
 11) Tweet 원문 : “We're being told we may have to stay at home for a few weeks. Many (like me) cannot work from home so possibly face time off unpaid. Many won't have to jobs to go back to. People are in debt up to their eyeballs, have big mortgage/rents to pay”, Mar. 11, 2020

<표 6> COVID-19 구간별 부정요인

구간	개수	부정요인	계수	SE ¹²⁾ 계수	t-값	p-값	VIF ¹³⁾
이전	1개	t22(원격직업)	-0.1365	0.0310	-4.40	0.0000	1.41
초기	9개	t22(원격직업)	-0.3076	0.0253	-12.18	0.000	1.15
		t18(거래/교환)	-0.2557	0.0243	-10.52	0.000	1.15
		t9(비전)	-0.2272	0.0207	-10.97	0.000	1.22
		t7(블로소득)	-0.1862	0.0136	-13.68	0.000	1.87
		t17(사이버보안)	-0.0916	0.0306	-2.99	0.003	1.09
		t35(팬테크)	-0.0892	0.0237	-3.77	0.000	1.22
		t47(기술)	-0.0822	0.0224	-3.67	0.000	1.23
		t4(재택사업)	-0.0752	0.0381	-1.98	0.048	1.06
		t3(기술직업)	-0.0625	0.0278	-2.25	0.024	1.11
심각기	14개	t18(거래/교환)	-0.0396	0.0098	-4.03	0.000	1.01
		t7(블로소득)	-0.0395	0.0051	-7.78	0.000	1.03
		t42(토론)	-0.0357	0.0101	-3.54	0.000	1.01
		t3(기술직업)	-0.0313	0.0106	-2.95	0.003	1.01
		t40(정책)	-0.0302	0.0091	-3.34	0.001	1.02
		t2(교육기술)	-0.0298	0.0098	-3.05	0.002	1.01
		t10(비즈니스 성장)	-0.0286	0.0100	-2.85	0.004	1.01
		t26(비즈니스 오너)	-0.0203	0.0116	-1.75	0.079	1.01
		t43(원격사무실)	-0.0192	0.0108	-1.78	0.075	1.01
		t13(셀프케어)	-0.0186	0.0100	-1.86	0.062	1.01
		t49(CIO)	-0.0185	0.0105	-1.76	0.078	1.01
		t1(자바개발자)	-0.0178	0.0105	-1.69	0.090	1.01
		t46(업적/성취)	-0.0163	0.0103	-1.59	0.112	1.01
		t34(S/W엔지니어)	-0.0093	0.0053	-1.74	0.081	1.03

주요, 전세계적으로 높은 역량이 있는 인력의 고용을 가능하게 한다¹⁴⁾와 같은 의견이 있다. 또한, AI/에자일 전문가, 클라우드 전환, VPN 과 같은 IT 관련된 토픽도 IT 인프라 지원이 필요한 비대면 근무 특성상 긍정적인 영향을 주고 있다.

반면, COVID-19 구간에 따라 긍정에서 부정으로 변경된 토픽으로는, 비즈니스 성장, 정책,

원격사무실 등이 있다. 팬데믹 선언 이후 대부분의 기업이 사전 준비 없이 비대면 근무를 시행하게 되었고, 기간 또한 장기화됨에 따라 비즈니스 성장에 대한 기대가 감소하였으며 정책에 대한 체계적인 대응이 미흡함을 알 수 있다. “기업은 갑자기 원격근무로 전환하였으나 전략적 준비는 없다¹⁵⁾”, “원격근무의 통제 부족은 업무 연속성, 보안, 효율성과 같은 비즈니스 향

12) SE : Standard Error, 표준오차

13) VIF : Variance Inflation Factor, 분산팽창요인

14) Tweet 원문 : “Remote work gives flexibility and independence, and allows us to hire top talent from across the globe”, Dec. 13, 2019

15) Tweet 원문 : “Organizations are suddenly remote, but didn't have a set of Remote Work strategy. Head of Remote will discuss how to implement processes of distributed working”, April 2, 2020

상에 큰 차질을 초래할 수 있다¹⁶⁾”는 의견을 예
 로 들 수 있다. 원격사무실의 경우, 재택근무로
 인해 가족간의 불편, 업무 집중의 어려움으로
 부정적인 의견으로 변경되었다. “재택근무자들
 은 오전 9시에서 오후 5시까지 호텔을 빌려 워
 크앤라이프 밸런스를 유지하려고 한다¹⁷⁾”는 의
 견이 그 예이다.

<표 7> 토픽별 감성 변화추이

구분	토픽	이전		초기		심각기	
t1	자바개발자	긍정	0.7796	긍정	0.4558	부정	-0.0178
t2	교육기술	긍정	0.7818	긍정	0.4479	부정	-0.02984
t3	기술직업	긍정	0.2501	부정	-0.0625	부정	-0.0313
t4	재택사업	긍정	0.0994	부정	-0.0752	-	-
t5	생산성	긍정	0.4009	긍정	0.1114	긍정	0.0104
t6	모바일 워크플로	긍정	0.3482	긍정	0.2052	-	-
t7	블로소득	긍정	0.0757	부정	-0.1862	부정	-0.03945
t8	재정적 자유	긍정	0.2748	긍정	0.1874	-	-
t9	비전	긍정	0.2072	부정	-0.2272	-	-
t10	비즈니스 성장	긍정	0.3722	긍정	0.0569	부정	-0.0286
t11	앱 플랫폼	긍정	0.3842	긍정	0.3992	-	-
t12	프리랜서	긍정	0.4899	긍정	0.6228	-	-
t13	셀프케어	긍정	0.445	긍정	0.0529	부정	-0.01862
t14	홈데코	긍정	0.635	긍정	0.177	-	-
t15	분석 연구원	긍정	0.5982	긍정	0.2822	-	-
t16	AI/에자일 전문가	긍정	0.2302	긍정	0.3382	-	-
t17	사이버 보안	긍정	0.3253	부정	-0.0916	-	-
t18	거래/교환	긍정	0.2281	부정	-0.2557	부정	-0.03956
t19	협업	긍정	0.5911	긍정	0.2876	-	-
t20	고립	긍정	0.2515	긍정	0.104	-	-
t21	디지털 전환	긍정	0.2532	-	-	-	-
t22	원격직업	부정	-0.1365	부정	-0.3076	-	-
t23	클라우드 전환	긍정	0.4493	긍정	0.3341	-	-
t24	IT 기술	긍정	0.372	긍정	0.1800	-	-
t25	인증	긍정	0.1974	긍정	0.3260	-	-
t26	비즈니스 오너	-	-	-	-	부정	-0.0203
t27	구직 검색	긍정	0.5213	긍정	0.1279	-	-
t28	미래업무	긍정	0.204	긍정	0.4021	-	-
t29	VR/IoT	-	-	긍정	0.1410	-	-
t30	하이브리드H/W	긍정	0.3269	긍정	0.1399	-	-
t31	퇴직	긍정	0.4407	긍정	0.0808	-	-
t32	주택/미니멀리즘	긍정	0.2235	긍정	0.2323	-	-
t33	유연근무	긍정	0.2000	긍정	0.2258	긍정	0.03018
t34	S/W 엔지니어	긍정	0.2833	긍정	0.1878	부정	-0.0093

16) Tweet 원문 : “The lack of control in Remote Work can lead to a multitude of setbacks in the development of your Business : continuity, security or efficiency, among others”, May 27, 2020
 17) Tweet 원문 : “Remote workers around the world are renting hotel rooms from 9 a.m. to 5 p.m. to try and regain a sense of work-life-balance”, June 22, 2020

t35	핀테크	긍정	0.3455	부정	-0.0892	-	-
t36	컨설턴트	긍정	0.2850	긍정	0.3311	긍정	0.01597
t37	조연자	긍정	0.3553	긍정	0.2428	-	-
t38	디지털 노마드	긍정	0.3838	긍정	0.2181	-	-
t39	VPN/웨비나	긍정	0.8183	긍정	0.6224	-	-
t40	정책	긍정	0.5538	긍정	0.2505	부정	-0.03022
t41	자동화	긍정	0.5982	-	-	-	-
t42	토론	긍정	0.3508	긍정	0.5546	부정	-0.0357
t43	원격 사무실	긍정	0.2301	-	-	부정	-0.0192
t44	리더십/기업가	긍정	0.4184	긍정	0.2623	-	-
t45	가상비서	긍정	0.4536	긍정	0.4192	-	-
t46	업적/성취	긍정	0.1199	-	-	부정	-0.0163
t47	기술	긍정	0.3627	부정	-0.0822	-	-
t48	협력업체	긍정	0.2017	-	-	-	-
t49	CIO	긍정	0.3547	긍정	0.3310	부정	-0.0185
t50	테라피/스킨케어	긍정	0.5399	긍정	0.0911	-	-

COVID-19 확산 방지를 위해 시행된 비대면 근무는 긍정적으로 유지되는 토픽도 있지만, 보안, 사무공간, 커뮤니케이션, 동기부여 등 토픽은 업무환경에 부정적인 영향을 미치고 있었다. 이는, 재택, 카페, 숙박시설과 같은 다양한 근무환경의 보안 취약점 대응 미흡, 원격 사무환경에서의 고객과의 관계 형성 어려움, 비대면 근무 시 상사와 동료 간 업무와 커뮤니케이션 부족, 업무시간과 개인시간의 구분 모호, 성과측정과 기한 기준 미흡 등이 원인으로 파악된다.

4.6 시사점

비대면 근무 관련 검출된 부정요인에 대해 기업에서는 다양한 정책을 통해 부정적 영향을 완화하고 근무환경을 개선할 수 있을 것으로 생각된다. 이러한 비대면 근무 관련 개선책으로 먼저 보안과 관련해서는, 재택근무 시 다양한 접속환경, 해커의 사이버 공격 급증에 대비하여 취약 PC에 대한 업무환경 접속차단, VDI 적용으로 Local PC 저장제한, 이상 상황 감지 시 자

동녹화 등의 보안체계 개선을 통해 부정요인을 완화할 수 있을 것으로 생각된다. 둘째, 사무공간과 관련해서는, 재택근무 시 가족들로 인한 업무집중의 어려움, 고객과의 관계 유지 어려움, 고정성 근무공간 고비용 발생에 대응하여, 본사 사무실의 고정성 좌석은 자율적으로 변경하여 사무공간을 축소하고, 거점별 보안체계가 갖춰진 공유사무실을 확보하여 대면 관계가 유지되도록 하는 등의 지원이 필요한 것으로 해석할 수 있다. 셋째, 커뮤니케이션과 관련해서는, 상사, 동료와의 일상적인 대화를 통한 비공식적 관계 구축 및 공식적 회의의 부족으로 인한 의사소통 단절의 문제개선을 위해, 소규모 일별 미팅을 통해 진행사항 및 정서적 관심사항을 공유하거나 가상사무공간을 연결함으로써 상호 근무상황에 대해 공유함으로써 문제를 개선할 수 있을 것으로 보인다. 마지막으로 동기부여와 관련해서, 재택근무의 반복되는 패턴으로 번아웃이 증가하고 성과수준 및 기한 가이드가 부족한 것에 대비하여, 재택근무에 특화된 유연시간제를 도입하여 개인에 맞는 근무시

간을 운영하도록 지원함으로써 문제를 개선할 수 있을 것이다.

COVID-19가 끝나는 시점인, 포스트 코로나에도 비대면 근무는 일정 부분 지속될 것으로 예상된다. 맥킨지에서는 “팬데믹 상황에서 재택근무가 성공적으로 자리 잡았지만 이러한 근무형태가 바로 뉴노멀이 되기는 어렵다. 개인과 소규모 팀의 생산성 향상, 비용 절감, 개인 유연성 향상 등 차원에서 하이브리드 모델을 효과를 얻을 것”(McKinsey Global Institute, MSI, 2021)으로 전망하였다. 강제정(2020)은 일과 삶의 균형을 위해 통근근무와 원격근무를 병행하여 대면접촉을 통한 사회적 유대감의 중요성을 강조하였다. 이와 같이 COVID-19 이후 비대면과 대면 근무가 병행된다면, 근무환경과 관련한 부정요인은 지속적으로 제기될 수 밖에 없다. 기업은 재택과 사무근무를 병행하기 위한 정책, 기준, 공간개선을 준비하고, 개인의 유연성 및 팀 내 협업을 유지하기 위한 업무문화를 구성하기 위해 근로자의 의견을 실시간으로 감지하고 반영할 수 있어야 한다. 비대면 환경으로 인한 비정형 데이터 증가로 감정분석의 중요성이 커지고 있다(이종화, 2020). 설문조사 또는 인터뷰 방식으로 문제점을 파악하고, 개선 방향을 제시하기에는 시간적, 물리적 한계가 존재하므로, 소셜 미디어 데이터를 활용하여 실시간 부정요인을 파악하는 본 연구방법은 향후에도 지속적으로 활용이 가능하다는 장점이 있다.

V. 결론

본 연구에서는 사회적으로 발생하는 다양한

이슈에 대해 즉각적으로 반응하는 소셜 미디어를 활용하여 비대면 근무와 관련한 일반인의 의견을 분석하였다. 2019년 7월부터 2020년 12월까지 18개월간의 트위터 데이터에 대해 COVID-19 구간에 따라 동적 토픽 모델링을 이용하여 토픽 변화를 분석하고, 감성 분석과 회귀분석을 통하여 부정적인 영향을 주는 요인을 검출하였다.

비대면 근무 관련 트윗 154만 건을 분석한 결과, COVID-19 이후 사용자 ID는 7.2배, 트윗 건수는 4.8배 증가하였다. 비대면 근무 관련 상위빈도 단어는 원격직업, 사이버보안, 기술직업, 생산성, 소프트웨어 순으로 나타났다. COVID-19 이후 증가한 단어는 락다운과 해고에 대한 우려와 업무복귀에 대한 기대감도 있다. 비즈니스 연속성 확보를 위한 가상미팅, 비즈니스 전환에 대한 대비책도 언급되었다. 사회적 활동이 감소함에 따라 수익창출, 체중감량, 뷰티, 여행 팁, 승진과 같은 단어의 빈도도 함께 감소하였다. 신규 발생한 단어는 새로운 표준을 의미하는 뉴노멀과 비대면 근무 도구인 줌 미팅, 구글 미트 등이 나타났다.

COVID-19 초기 구간내 정점(Peak) 토픽으로는, 비대면 근무 IT 인프라 지원을 위한 CIO, 비즈니스 성장에 대한 우려, 주택 및 미니멀리즘이 나타났다. 또한, 비대면 근무가 장기화됨에 따라 COVID-19 심각기 구간 내 정점 토픽은, 재정적인 자유에 대한 관심 증가, 정책에 대한 확대 필요성 및 비대면 근무를 위한 가상지원 등이 있었다. COVID-19 구간별 감성 분석을 비교한 결과, 부정의견은 이전 34%에서 초기 43%로 증가하였다가 심각기에 36%로 안정화되었다. 회귀분석을 통해 감정 점수에 영향을

주는 토픽을 분석한 결과, COVID-19 이후 부정요인으로 비즈니스 성장, 사이버보안, 정책, 원격사무실 등이 나타났다. COVID-19 팬데믹 선언 이후 비대면 근무가 장기화됨에 따라 비즈니스 성장에 대한 기대는 감소하였고, 보안 및 정책에 대한 체계적인 대응 미흡하여 부정적인 영향을 주고 있다. COVID-19 이후에도 재택근무에 대한 부정요인은 지속적으로 제기될 수 있으므로 기업에서는 지속적인 부정요인의 분석 및 대책 마련이 필요하다고 할 수 있다.

본 연구에서는 많은 시간과 인력이 동반되는 기존의 설문조사 방식을 토픽 모델링과 감성 분석을 통한 정량적 분석으로 전환하는 방법론을 채택하였다. 토픽이 감성에 미치는 영향을 보기 위해 토픽을 독립변수로 하여 회귀분석을 실시함으로써, 단지 부정문서와 긍정문서의 토픽을 비교하는 것보다 체계적이고 정량적인 분석이 가능한 방법을 제시했다는 점에서 학술적 기여가 있는 것으로 판단된다. 또한 COVID-19 이후 비대면 근무에 대한 부정요인을 분석하기 위한 방법과 결과 및 시사점을 제시함으로써, 기업들이 근무체계 개선에 활용할 수 있도록 했다는 점에서 실무적인 기여가 있다고 생각된다.

본 연구의 한계는 산업군별, 기업유형별 비대면 근무체계 준비 및 개선 수준과 무관하게 일반화된 트위터 데이터를 기반으로 분석함으로써, 서로 다른 환경에 있는 기업의 특수한 상황을 고려하지 못한 데 있다. 이를 보완하기 위해서는 일반적인 트윗 대신 각 기업의 사내 포털 사이트의 데이터를 대상으로 하거나, 그 외 기업 혹은 산업과 관련있는 언론사 기사 및 전문 리서치기관의 자료 등을 활용함으로써, 산업과 기업

에 보다 특화된 분석을 하고 비대면 근무의 개선점을 도출할 수 있을 것으로 기대된다.

참고문헌

- 강재정, “원격근무환경에서 일과 삶의 균형방안: 인과지도, 사고프로세스와 트리즈의 적용”, 한국정보시스템학회, 정보시스템연구, 29권 3호, 2020, pp. 179-195
- 강찬희, “동적 토픽 모델링과 의미연결망을 통한 코로나19 이슈분석”, 국내 석사학위논문, 성균관대학교 일반대학원, 서울, 2021
- 고은지, 최선영, “트윗의 타임 시퀀스를 활용한 DTM 분석 : 2019 남북미정상회담 이벤트를 중심으로”, 한국정보통신학회 논문지, 25(2), 2021, pp. 308-313
- 박일경, “재택형 유연근무제의 실태에 관한 연구”, 2019, 국내 석사학위논문, 인천대학교, 인천
- 서창적, 김효정, “텔파이 기법을 이용한 비대면 서비스 재택근무 적합요인의 도출”, 서비스경영학회지 14.1, 2013, pp. 207-227
- 우창우, 이종연, “다이나믹 토픽 모델을 활용한 D(Data) · N(Network) · A(A.I) 중심의 연구동향 분석”, 한국융합학회논문지, 11(9), 2020, pp. 21-29
- 이종화, “SNS대상의 지능형 자연어 수집, 처리 시스템 구현을 통한 한국형 감성사전 구축에 관한 연구”, 한국정보시스템학회, 정보시스템연구 29권 3호, 2020,

- pp. 237-251
- 이주양, “소셜 미디어 감성 분석을 활용한 항공사 고객만족도 연구 : 한, 영, 일, 중 언 어권을 중심으로”, *관광경영연구*, 24(6), 2020, pp. 527-546
- 이진영, “스마트워크가 자기효능감과 업무창의성 및 조직성과에 미치는 구조관계 연구”, 국내 박사학위논문 경희대학교 대학원, 서울, 2020
- 이혜진, 최자영, “텍스트 마이닝 기법을 활용한 편의점 이용 고객의 트위터 리뷰에 대한 감성 분석”, *글로벌 경영학회지*, 2019, pp. 143-164
- 정용주, 김진수, “재택근무 제도강화와 테크노스트레스가 업무생산성에 미치는 영향에 관한 연구 : 행동통제와 기술준비도의 조절효과를 중심으로”, *JITAM*, 27(4), 2020, pp. 63-83
- 차영란, “코로나 이전과 이후의 4차 산업혁명과 광고의 뉴스기사 분석 : LDA와 Word2vec을 중심으로”, *한국콘텐츠학회 논문지 제21권 제9호*, 2021, pp. 149- 163
- 최홍열, 박은경, “소셜 미디어 빅데이터 분석을 이용한 나홀로 여행 트렌드 분석 : 제주도를 중심으로”, *관광경영연구*, 23(1), 2019, pp. 45-64
- Alec Go, Richa Bhayani, Lei Huang, “Twitter Sentiment Classification Using Distant Supervision”, *CS224N Project Report, Stanford*, 2009, pp. 12
- Andrea Alexander, Rich Cracknell, Aaron De Smet, Meredith Langstaff, Mihir Mysore, and Dan Ravid, “What Executives Are Saying about the Future of Hybrid Work”, *McKinsey Insights*, 2021
- Andrew N. O'Connor, “Probability Distributions Used in Reliability Engineering”, *Center for Risk and Reliability*, 2011
- Bondarenko Olena M., Tardaskina Tetiana M., Bogatyryova Lyudmyla D., “Researching the Development of Remote Work in the Context of the COVID-19 Pandemic”, *BIZNES INFORM*, Vol. 2, no. 517, 2021, pp. 190-197
- Bo Pang, Lillian Lee, “Opinion Mining and Sentiment Analysis”, Boston : *Now Publishers*, 2008, pp. 1-135
- Changqin Quan, Fuji Ren, “Unsupervised Product Feature Extraction for Feature-Oriented Opinion Determination”, *Information Sciences*, Vol. 272, 2014, pp. 16-28
- David M. Blei, “Latent Dirichlet Allocation”, *Journal of machine learning research*, ISSN : 1532-4435, 2003, pp. 993-1022
- Farah Shahid, Aneela Zameer, Muhammad Muneeb, “Predictions for COVID-19 with Deep Learning Models of LSTM, GRU and Bi-LSTM”, *Chaos, Solitons & Fractals*, Vol. 140, 2020
- Fu Xianghua, Liu Guo, Guo Yanyan, Wang Zhiqiang, “Multi-Aspect Sentiment Analysis for Chinese Online Social

- Reviews Based on Topic Modeling and HowNet Lexicon”, *Knowledge-Based Systems*, Vol. 37, 2013, pp. 186-195
- Jeremy Barnes, Roman Klinger, Sabine Schulte Im Walde, “Assessing State-of-the-Art Sentiment Models on State-of-the-Art Sentiment Datasets”, *In Proceedings of WASSA*, 2017, pp. 2-12
- J. Grimmer, “A Bayesian Hierarchical Topic Model for Political Texts : Measuring Expressed Agendas in Senate Press Releases”, *Political Analysis*, 18, 2010, pp. 1-35
- Jian Jin, Ping Ji, Rui Gu, “Identifying Comparative Customer Requirements from Product Online Reviews for Competitor Analysis”, *Engineering Applications of Artificial Intelligence*, Vol. 49, 2015, pp. 61-73
- Kasper Christensen, Kristian Hovde Liland, Knut Kvaal, Einar Risvik, Alessandra Biancolillo, Joachim Scholderer, Sladjana Nørskov, Tormod Næs, “Mining Online Community Data : The Nature of Ideas in Online Communities”, *Food Quality and Preference*, Vol. 62, 2017, pp. 246-256
- Mathieu Cliche, “BB_twtr at SemEval-2017 Task 4: Twitter Sentiment Analysis with CNNs and LSTMs”, *Proceedings of SemEval*, 2017, pp. 8
- Mohamed M. Mostafa, “More than Words : Social Networks’ Text Mining for Consumer Brand Sentiments”, *Expert Systems with Applications*, Vol. 40, Issue 10, 2013, pp. 4241-4251
- Neal Schmitt, “Method Bias: The Importance of Theory and Measurement”, *Journal of Organizational Behavior*, Vol. 15, no. 5, Wiley, 1994, pp. 393-398
- Nikolay Archak, “Deriving the Pricing Power of Product Features by Mining Consumer Reviews”, *Management Science*, 57-8, 2011, pp. 1485-1509
- Nut Limsopatham and Nigel Collier, “Bidirectional LSTM for Named Entity Recognition in Twitter Messages”, *The COLING 2016 Organizing Committee*, 2016, pp. 145-152
- Paul C. Tetlock, “Giving Content to Investor Sentiment: The Role of Media in the Stock Market”, *The Journal of Finance*, Vol. 62, Issue 3, 2007, pp. 1139-1168
- Priyanka Vedurumudi, “Twitter Sentiment Analysis using Deep Learning”, *I-Manager's Journal on Computer Science*, 2021, pp. 1
- Teresa Galanti, Gloria Guidetti, Elisabetta Mazzei, Salvatore Zappalà and Ferdinando Toscano, “Work From Home During the COVID-19 Outbreak: The Impact on Employees' Remote Work Productivity, Engagement, and Stress”, *J Occup Environ Med* ; 63(7), 2021, pp. 426-432
- Tim O'Reilly, “What Is Web 2.0: Design

Patterns And Business Models For The Next Generation Of Software”, University Library of Munich, Germany, *MPRA Paper*, 2007

Xuefeng Wang, Xuemei Yang, Xinglu Wang, Mengge Xia and Jiayun Wang, “Evaluating the Competitiveness of Enterprise’s Technology based on LDA Topic Model”, *Technology Analysis & Strategic Management*, 32:2, 2020, pp. 208-222

Ya-Han Hu, Yen-Liang Chen and Hui-Ling Chou, “Opinion Mining from Online Hotel Reviews - A Text Summarization Approach”, *Information Processing & Management*, Vol. 53, Issue 2, 2017, pp. 436-449

Yee Whye Teh, “A Hierarchical Bayesian Language Model Based On Pitman-Yor Processes”, *Association for Computational Linguistics*, 2006, pp. 985-992

Yikai Yan and Chaoyue Niu, Yucheng Ding, Zhenzhe Zheng, Fan Wu, Guihai Chen, Shaojie Tang, Zhihua Wu, “Distributed Non-Convex Optimization with Sublinear Speedup under Intermittent Client Availability”, arXiv:2002.07399, 2020

Zeng, Jianping, Duan, Jiangjiao, Wang, Wei and Wu, Chengrong, “Semantic Multi-Grain Mixture Topic Model for Text Analysis”, *Expert Syst. Appl.* 38, 2011, pp. 3574-3579

이 선 민 (Lee, Sun Min)



McGill 대학교 전자공학 학사와 Toronto 대학교 경영 학과 석사를 취득하였다. 현재 연세대학교 산업공학과 박사 과정에 있으며, 삼성전자에 재직 중이다. 주요 관심분야는 비즈니스 애널리틱스, 딥러닝, 텍스트 마이닝 등이다.

천 세 진 (Chun, Se Jin)



서울시립대 컴퓨터과학 학사와 연세대학교 석사와 박사 학위를 취득하고, 미 표준기술 연구원(NIST)에서 박사후 과정을 수행하였다. 현재 동아대학교 컴퓨터 AI공학부 조교수로 재직하고 있으며, 주요 관심분야는 지식 그래프 기반 추론 등이다.

박 상 언 (Park, Sang Un)



한국과학기술원 전산학사, 경영공학 석사와 박사학위를 취득하였다. 현재 경기대학교 경영정보전공 교수로 재직하고 있으며, 주요 관심분야는 인공지능, 머신러닝, 딥러닝, 텍스트 마이닝 등이다.

이 태 욱 (Lee, Tae Wook)



연세대학교 산업공학학사를 취득하였고, 현재 연세대학교 산업공학과 석사과정에 있다. 주요 관심분야는 빅데이터 분석, 머신러닝, AI, 자연어 처리, 추천시스템 등이다.

김 우 주 (Kim, Woo Ju)



연세대학교 BBA과정 학사를 취득하고, KAIST 경영과학 석사, 박사를 취득하였으며, 현재 연세대학교 산업공학과 교수로 재직 중이다. 관심분야는 딥러닝, 온톨로지, 텍스트 마이닝, 이상 탐지, 이미지 분석 등이다.

<Abstract>

Detection of Complaints of Non-Face-to-Face Work before and during COVID-19 by Using Topic Modeling and Sentiment Analysis

Lee, Sun Min · Chun, Se Jin · Park, Sang Un · Lee, Tae Wook · Kim, Woo Ju

Purpose

The purpose of this study is to analyze the sentiment responses of the general public to non-face-to-face work using text mining methodology. As the number of non-face-to-face complaints is increasing over time, it is difficult to review and analyze in traditional methods such as surveys, and there is a limit to reflect real-time issues.

Approach

This study has proposed a method of the research model, first by collecting and cleansing the data related to non-face-to-face work among tweets posted on Twitter. Second, topics and keywords are extracted from tweets using LDA(Latent Dirichlet Allocation), a topic modeling technique, and changes for each section are analyzed through DTM(Dynamic Topic Modeling). Third, the complaints of non-face-to-face work are analyzed through the classification of positive and negative polarity in the COVID-19 section.

Findings

As a result of analyzing 1.54 million tweets related to non-face-to-face work, the number of IDs using non-face-to-face work-related words increased 7.2 times and the number of tweets increased 4.8 times after COVID-19. The top frequently used words related to non-face-to-face work appeared in the order of remote jobs, cybersecurity, technical jobs, productivity, and software. The words that have increased after the COVID-19 were concerned about lockdown and dismissal, and business transformation and also mentioned as to secure business continuity and virtual workplace. New Normal was newly mentioned as a new standard. Negative opinions found to be increased in the early stages of COVID-19 from 34% to 43%, and then stabilized again to 36%

————— 동적 토픽 모델링과 감성 분석을 이용한 COVID-19 구간별 비대면 근무 부정요인 검출에 관한 연구

through non-face-to-face work sentiment analysis. The complaints were, policies such as strengthening cybersecurity, activating communication to improve work productivity, and diversifying work spaces.

Keyword: Dynamic Topic Modeling, Remote Work, Sentiment Analysis, Twitter Analysis,
Work from Home

* 이 논문은 2021년 12월 1일 접수, 2021년 12월 10일 1차 심사, 2021년 12월 20일 게재 확정되었습니다.