

경영전문석사학위 논문

기업 리뷰 데이터를 활용한

ESG 경영 가치 연구

- 기업의 사회 책임 경영을 중심으로 -

2023년 2월

서울과학종합대학원대학교

이 승 엽

경영전문석사학위 논문

기업 리뷰 데이터를 활용한

ESG 경영 가치 연구

- 기업의 사회 책임 경영을 중심으로 -

2023년 2월

서울과학종합대학원대학교

이 승 엽

초 록

ESG 경영이 비재무적 가치로서 기업 경영에 주요한 전략으로 자리잡게 되면서, 기업의 사회 책임 경영에 대한 관심 또한 높아졌다. 그에 따라, 기업의 내부 관계자인 임직원에게 긍정적인 직원 경험을 설계하여 직무 안정감을 부여하고, 이탈을 방지하기 위한 인적 자원 관리 또한 중요해졌다.

기존 ESG 등급에 대한 국내의 연구는 비재무적 요소를 통해 재무적 가치를 추산하고자 하는 연구가 주를 이루었으며, 주로 설문조사를 활용하여 표본의 수가 많지 않아 대표성을 가지기가 어렵다는 한계점이 있었다.

따라서 본 연구는 ESG 등급 중에서 사회(S) 등급을 그룹별로 나누고, 대퇴사·대이직 시대에 그 가치가 더욱 높아지고 있는 잡플래닛 기업 리뷰 데이터 분석을 통해, 추출되는 토픽의 차이를 살펴봄으로써, ESG 경영 수준에 따른 기업의 사회 책임 경영 전략에 대한 실무적인 시사점을 제시하는 것에 그 목적이 있다.

연구 대상은 2022 년 9 월 기준 KOSPI 시가 총액 상위 404 개 기업이며, 한국 ESG 기준원에서 평가한 2021 년 사회(S) 등급과 해당 기업의 잡플래닛 리뷰 데이터 104,452 개를 텍스트 마이닝하고, 토큰화 및 LDA 토픽모델링 기법을 활용하여 분석하였다.

연구 결과, 첫째, 장점 리뷰에서는 일과 삶의 균형을 뜻하는 ‘워라밸’ 토픽이 상, 중 그룹에서만 언급되었고 하 그룹에서는

언급되지 않았다. 또한, 하 그룹에서는 수평적인 조직문화에 대한 언급이 없었다. 마지막으로, 하 그룹에서는 현장 근무에 대한 조건들이 자세하게 나타나고 있음을 확인하였다. 이를 통해, 하 그룹에 속하는 기업에서는 근로자의 일과 삶의 균형을 보장할 수 있는 가족친화 경영 및 조직문화 개선 활동, 안전보건 거버넌스 체계 구축 등이 필요함을 시사한다.

둘째, 단점 리뷰에서는 상, 중 그룹에서는 ‘문서 업무’ 라는 토픽이 도출되었다. 또한, 하 그룹에서는 ‘임원의 의사결정’ 이라는 토픽이 주요 토픽으로 도출되었다. 이를 통해, 상, 중 그룹에서는 스마트 워크 환경 구축을 통해 업무 효율성을 제고하고, 하 그룹에서는 임원의 의사결정력 제고 및 적절한 권한 부여가 필요하다.

마지막으로, 경영진에 바라는 점 리뷰에서는 상 그룹에서만 ‘비정규직’ 이라는 토픽이 도출되었는데, 비정규직의 정규직 전환 및 처우 개선에 대한 요구가 나타났다. 또한, 중 그룹에서만 ‘인사제도’ 토픽이 도출되었다. 마지막으로 하 그룹에서만 ‘경영진의 인식 개선’ 이라는 토픽이 도출되었다. 이를 통해, 상 그룹에서는 비정규직 처우 개선을 통한 노사관계 개선이 필요하며, 중 그룹에서는 모든 구성원들에게 공정한 기회 제공 및 평가를 시행하는 인사제도 마련이 필요하다. 마지막으로, 하 그룹에서는 직원들을 회사의 미래라고 인식하고 성장시키기 위한 경영진의 인식 개선이 필요함을 시사하였다.

본 연구는 온라인 기업 리뷰 분석을 통해 긍정적인 직원 경험을 설계하고, 내부 직원의 이탈을 방지하기 위한 실무적인 시사점을

제시하는 HR Analytics 를 수행하였으며, 기업의 ESG 경영 수준별로 차별점을 두었다는 점에서 새로운 연구방향을 제시하였다.

목 차

제 I 장 서론	1
제 1절 연구배경.....	1
제 2절 연구목적 및 구성	3
제 II 장 이론적 배경 및 선행연구 탐색	4
제 1절 이론적 배경.....	4
제 2절 선행연구 분석.....	11
제 III 장 본론	17
제 1절 연구문제 및 연구설계	17
제 2절 연구대상 및 수집	18
제 3절 연구방법	24
제 4절 토픽모델링 결과	33
제 IV 장 결론	50
제 1절 연구 결과	50
제 2절 연구의 시사점	52
제 3절 연구의 한계점 및 제언	53
참고문헌	54

표 목 차

<표1> Examples of ESG Issues	5
<표2> 한국ESG기준원 평가 모형 체계	6
<표3> 한국ESG기준원 ESG 등급 명칭 및 의미	8
<표4> 불용어 목록	25
<표5> 각 리뷰별 최적의 토픽 갯수 및 일관성 지표	27
<표6> 전체 '장점' LDA 토픽모델링 결과	35
<표7> 상 '장점' LDA 토픽모델링 결과	36
<표8> 중 '장점' LDA 토픽모델링 결과	37
<표9> 하 '장점' LDA 토픽모델링 결과	38
<표10> 전체 '단점' 리뷰 LDA 토픽모델링 결과	40
<표11> 상 '단점' LDA 토픽모델링 결과	41
<표12> 중 '단점' LDA 토픽모델링 결과	42
<표13> 하 '단점' LDA 토픽모델링 결과	43
<표14> 전체 '경영진에 바라는 점' LDA 토픽모델링 결과	46
<표15> 상 '경영진에 바라는 점' LDA 토픽모델링 결과	47
<표16> 중 '경영진에 바라는 점' LDA 토픽모델링 결과	48
<표17> 하 '경영진에 바라는 점' LDA 토픽모델링 결과	49

그림 목 차

<그림1> LDA Plate Notation	10
<그림2> 연구모형	18
<그림3> 잡플래닛 기업 리뷰 데이터 구성	19
<그림4> S 등급별 기업 분포	20
<그림5> 전/현직자별 잡플래닛 기업 리뷰 갯수	21
<그림6> 직무별 잡플래닛 기업 리뷰 갯수	22
<그림7> S 등급별 잡플래닛 기업 리뷰 갯수	23
<그림8> 장점(전체) LDA 토픽모델링 시각화	31
<그림9> 단점(전체) LDA 토픽모델링 시각화	32
<그림10> 경영진에 바라는 점(전체) LDA 토픽모델링 시각화	32

제 I 장 서론

제 1 절 연구 배경

미국 노동통계국(U.S. Bureau of Labor Statistics) 자료(Job Openings and Labor Turnover Survey, 2021)에 따르면, 2021 년 11 월 한 달 동안의 미국의 자발적 퇴사자는 약 453 만 명으로 보고되었다. 이는 관측을 시작한 2000 년 12 월 이후 가장 높은 수치인데, Anthony Klotz(2022)는 이러한 현상을 "대퇴사(Great Resignation) 시대가 도래한 것"이라고 설명한다. COVID-19 로 촉발된 번아웃(Burnout) 현상을 겪으며 많은 근로자들이 직무만족에 대한 가치를 다시금 생각하게 되었으며, 더 나은 보상체계와 직무만족을 이루기 위해 회사를 떠나고 있다.

대퇴사 현상이 비단 미국에서만 나타나는 국한된 현상은 아니다. 고용노동부에서 발표한 2021 년 12 월 사업체 노동력조사 결과에 따르면, 2021 년 누계 국내 이직자 수는 1,105 만 7,000 명으로, 이는 10 년 전과 비교하면 약 두 배 가량 증가한 수치다. 바야흐로 많은 근로자들이 더 나은 처우와 근무환경, 직무만족을 위해 끊임없이 이직을 시도하는 전 세계적인 대이직 시대가 열린 것이다. 기업에서는 근로자의 이탈을 막기 위해 COVID-19 를 대비한 스마트워크 환경을 구축하고, 보상 체계를 강화하는 등 다양한 교육지책들을 강구하고 있지만, 시대의 흐름을 거스르기란 쉽지 않아 보인다.

Josh Plaskoff(2017)는 직원경험(Employee Experience)을 “입사부터 퇴사에 이르기까지 직장 내 여정 전체에 걸쳐 회사와 자신의 관계에 대한 전체적인 인식”이라고 정의한다. 클래스도어와 잡플래닛으로 대표되는 온라인 기업 리뷰는 이러한 직원 경험을 익명 작성을 기반으로 불특정 다수와 공유하는 수단이다. 이러한 온라인 기업 리뷰는 대퇴사 시대가 도래함에 따라 이직을

원하는 근로자들에게 더욱 가치 있는 데이터가 되었다. 온라인 기업 리뷰는 데이터를 쉽게 수집할 수 있으며, 항목별 평점 데이터와 함께 비정형 텍스트 데이터로 구성되어 있기 때문에 다양한 분석이 가능하여 최근 많은 연구가 이루어지고 있는 주제이기도 하다.

동시에 ESG 가 비재무적 가치로서 기업 경영에 주요한 전략으로 자리잡게 됨에 따라, ESG 는 최근 학술 연구의 주요한 주제로 부각되었다. ESG 란 환경(Environment), 사회(Social), 지배구조(Government)의 약자로, ESG 의 두 번째 요소인 Social 의 내용을 살펴보면 인적자원(Human Resources) 분야에서 다루고 있는 이슈들과 상당히 밀접하게 연관되어 있다(이재진, 2022). 유엔 책임투자원칙에서 제시하는 ESG 의 내용을 보면, Social 의 Human Right, Working Condition, Employee Relation 등의 내용은 HR 의 영역으로 포함시킬 수 있다. 이렇듯, 기업의 ESG 경영은 인적자원 운영(Human Resources Management)의 관점으로도 이해할 수 있다.

그러나, 국내에서 기업의 ESG 경영에서도 사회(S) 분야에 대한 연구는 주로 기업의 사회적 책임(Coporate Social Reponsibility, CSR) 혹은 사회 책임 투자(Social Responsibility Investment, SRI)의 차원에서 주로 논의되어 왔으며, HR 의 관점으로 진행된 연구는 아직 미진한 편이다(박재현, 한향원 & 김나라, 2022). 따라서 본 연구는 대퇴사 시대를 맞아, 기업의 ESG 등급 중에서도 사회(S) 등급을 구간별로 구분하여 각각의 근로자들의 직무만족, 직무불만족 요소와 경영진에 바라는 점을 기업 리뷰 데이터 분석을 통해 살펴보았다. 또한 직원경험을 긍정적으로 극대화하고, 근로자들의 이탈을 방지하기 위해서 기업에서는 어떠한 전략적인 경영을 해야 하는지 ESG 경영의 사회(S)와 연계한 HR Analytics 관점에서 실무적인 시사점을 제시하고자 하였다.

제 2 절 연구목적 및 구성

본 연구에서는 한국 ESG 기준원에서 평가한 기업의 2021 년 사회(S) 등급을 S, A+, A 는 상, B+, B 는 중, C 와 D 는 하로 분류하고, 각 그룹별로 해당 기업의 잡플래닛 리뷰 데이터를 Python 의 BeautifulSoup 패키지를 활용한 크롤링을 통해 수집하였다. 수집한 데이터는 기업의 장, 단점 및 경영진에 바라는 점으로 나눈 뒤 일관성(Coherence) 지표를 통해 최적의 토픽 갯수를 산정하고, LDA 토픽모델링을 통해 상, 중, 하 그룹 간에 어떠한 토픽들이 주로 언급되는지 살펴보고, 기업의 ESG 경영에서도 사회(S)적 요소에 초점을 맞추어 근로자의 이탈을 방지하기 위한 기업 경영의 시사점을 제시하고자 하였다.

본 연구의 구성은 다음과 같다. 2 장에서는 지속가능경영과 ESG 등급 및 데이터 분석에 활용되는 LDA 토픽모델링에 대해 설명하고, 선행연구를 살펴본다. 3 장에서는 데이터 수집과 분석에 대해 자세히 설명한다. 4 장에서는 연구 분석 및 결과에 대해 살펴보고, 마지막 5 장에서는 본 연구의 시사점 및 한계에 대해 서술한다.

제 II장 이론적 배경 및 선행연구 탐색

제 1 절 이론적 배경

본 연구에서는 한국ESG기준원에서 평가한 기업의 ESG 등급 중에서도 기업의 사회(Social) 등급을 구간별로 나누고, 구간별로 분리하여 잡플래닛 기업리뷰 데이터를 수집하여 LDA 토픽모델링을 통해 추출되는 토픽 간의 차이를 확인하였다. 전체 데이터와 대비하여 추출되는 토픽이 다르다는 것은 해당 그룹에서 특히 더욱 많이 언급되는 것으로, 이는 기업 내부 임직원의 이탈 방지와 긍정적인 직원경험 설계를 위해 우선적으로 점검해야 하는 내용이다. 따라서, 본 장에서는 ESG 경영에 대한 이론과, ESG 경영을 평가하는 척도인 ESG 등급에 대해 알아보고, 자연어 처리 기법인 토픽 모델링 기법 중 본 연구에서 활용할 잠재 디리클레 할당(LDA) 기법에 대해 정리한다.

(1) 지속가능 경영과 ESG(Environment, Social, Government) 등급

ESG란, 환경(Environment), 사회(Social), 지배구조(Governance)의 앞 글자로 구성된 단어로, 2006년 유엔 책임투자원칙(UN PRI : UN Principles for Responsible Investment)이 출범하면서 ESG 구성 요소가 처음으로 제시되었다. 아래 <표 1>은 UN PRI에서 제시하는 ESG 구성 요소이다. ESG는 기업의 재무적인 측면 외 지속가능경영을 판단하는 비재무적 기준으로 널리 활용되고 있다. 국내에서는 비교적 최근인 2010년대에 들어서 지속가능경영이라는 개념이 기업 경영의 화두가 된 가운데, 그 중요성은 점차 강조되고 있다. 금융위원회(2021)에 따르면, 오는 2025년부터는 자산 2조원 이상 등 일정 규모 이상의 코스피 상장사는 환경(E)와 사회(S) 정보를 의무적으로 공시하여야

하며, 2030년에는 모든 코스피 상장사로 그 범위가 확대된다. 또한 지배구조(G)는 2026년부터 모든 코스피 상장사가 의무적으로 공시하여야 하는 등 ESG는 기업 경영의 새로운 패러다임으로 그 가치를 점차 더해가고 있다.

<표 1> Examples of ESG Issues

환경(Environment)	사회(Social)	지배구조(Governance)
• Climate Change	• Human Rights	• Bribery and Corruption
• Resource Depletion	• Modern Slavery	• Executive Pay
• Waste	• Child Labour	• Board Diversity and Structure
• Pollution	• Working Conditions	• Political Lobbying and Donations
• Deforestation	• Employee Relations	• Tax Strategy

자료: UN Principles for Responsible Investment

기업의 ESG 경영을 측정하는 도구로는 다양한 기관의 평가지표가 활용되는데, 국내에서는 주로 한국ESG기준원(Korea Corporate Governance Service, KCGS)에서 평가하는 등급을 준용한다. 한국ESG기준원은 2002년 설립된 한국거래소(Korea Exchange, KRX) 산하 비영리단체로, 2011년부터 국내 유가증권 상장사에 대하여 지속가능경영 수준 평가를 진행해왔으며, 2022년에는 11월 기준 국내 유가증권 상장사 1,005개에 대한 평가를 진행하였다. ESG 평가는 ISO26000, OECD 기업지배구조 원칙 등 국제기준과 국내 법제 및 경영환경을 반영하여 개발한 독자적 평가모형을 따른다(한국ESG기준원, 2021). 아래 <표 2>는 한국ESG기준원이 제시하는 평가 모형 체계이다.

<표 2> 한국ESG기준원 평가 모형 체계

평가 영역	평가모형 모델	KCGS 모범기준	법·제도·규범
환경	<ul style="list-style-type: none"> ○ 경영 프로세스 모델 - 업종에 따른 환경위험관리 및 성과평가 고려 - 산업별 환경 민감도를 상·중·하로 구분하여 구성 - 환경경영 체계, 성과, 이해관계자 대응 수준 판단 	<ul style="list-style-type: none"> • 리더십과 거버넌스 • 위험관리 • 운영 및 성과 • 이해관계자 소통 	<ul style="list-style-type: none"> • 저탄소 녹색성장 기본법 • 화학물질의 등록 및 평가 등에 관한 법률 • ISO 14001 및 환경표지·환경인증제도 • 기후관련 재무정보공개 (TCFD) 등
사회	<ul style="list-style-type: none"> ○ 이해관계자 모델 - 기업과 직·간접적 이해관계를 형성한 대상에 중점 - 산업분류기준인 WICS 중분류로 분리하여 산업별로 이해관계자에게 중대한 사회책임경영 이슈 고려 	<ul style="list-style-type: none"> • 리더십과 거버넌스 • 비재무 위험 관리 • 운영 및 성과 • 이해관계자 소통 	<ul style="list-style-type: none"> • 독점규제 및 공정거래에 관한 법률 • 사회적 책임 세계 표준 (ISO 26000) • 국제노동기구 핵심 협약 • UN 국제 인권 선언 및 OECD 인권 실사 지침 등

지배 구조	<ul style="list-style-type: none"> ○ 통제기능 모델 - 지배구조가 작동하기 위한 주요 장치별 분류 - 일반 상장사 및 금융회사 특화 지배구조 요건 고려 - 자산 규모 및 이사회 내 주요 위원회 설치 여부, 세부 금융업권 특성 고려 	<ul style="list-style-type: none"> • 이사회 리더십 • 주주권 보호 • 감사 • 주주 및 이해관계자 소통 	<ul style="list-style-type: none"> • 상법 및 금융회사의 지배구조에 관한 법률 • OECD 기업지배구조 원칙 • 영국 FRC 및 일본 JPX 기업지배구조 코드 • ICGN 글로벌 지배구조 원칙 • 바젤은행감독위원회 은행 지배구조 원칙 등
-------	--	--	---

자료 : 한국ESG기준원(KCGS)

ESG 등급 평가는 기업의 공시자료와 기관, 미디어 자료를 기본 수집하여 활용하고, 18개 대분류의 총 265개 평가항목으로 기본평가를 진행한다. 또한 58개의 핵심 평가항목 심화 평가를 통해 기업의 ESG 등급을 평가하게 된다 (한국ESG기준원, 2021). 기업의 ESG 평가등급은 각 항목별로 S, A+, A, B+, B, C, D 총 7개 등급으로 분류되는데, 이 등급은 한국거래소(KRX)와 한국ESG기준원 홈페이지에 그 결과를 공시함으로써 개인, 기관의 ESG 투자에 대한 참고 자료로 활용되고 있다. ESG 등급 명칭 및 의미는 아래 <표 3>과 같다.

<표 3> 한국ESG기준원 ESG 등급 명칭 및 의미

평가 등급	의 미
S	지배구조, 사회, 환경 모범기준이 제시한 지속가능경영 체계를 매우 충실히 갖추고 있으며, 비재무적 리스크로 인한 주주가치의 훼손의 여지가 매우 적음
A+	지배구조, 사회, 환경 모범기준이 제시한 지속가능경영 체계를 충실히 갖추고 있으며, 비재무적 리스크로 인한 주주가치의 훼손의 여지가 상당히 적음
A	지배구조, 사회, 환경 모범기준이 제시한 지속가능경영 체계를 적절히 갖추고 있으며, 비재무적 리스크로 인한 주주가치의 훼손의 여지가 적음
B+	지배구조, 사회, 환경 모범기준이 제시한 지속가능경영 체계를 적절히 갖추기 위한 노력이 다소 필요하며, 비재무적 리스크로 인한 주주가치의 훼손의 여지가 다소 있음
B	지배구조, 사회, 환경 모범기준이 제시한 지속가능경영 체계를 적절히 갖추기 위한 노력이 다소 필요하며, 비재무적 리스크로 인한 주주가치의 훼손의 여지가 있음
C	지배구조, 사회, 환경 모범기준이 제시한 지속가능경영 체계를 적절히 갖추기 위한 노력이 절대적으로 필요하며, 비재무적 리스크로 인한 주주가치의 훼손의 여지가 큼
D	지배구조, 사회, 환경 모범기준이 제시한 지속가능경영 체계를 거의 갖추지 못하여 비재무적 리스크로 인한 주주가치의 훼손이 우려됨

자료 : 한국ESG기준원(KCGS)

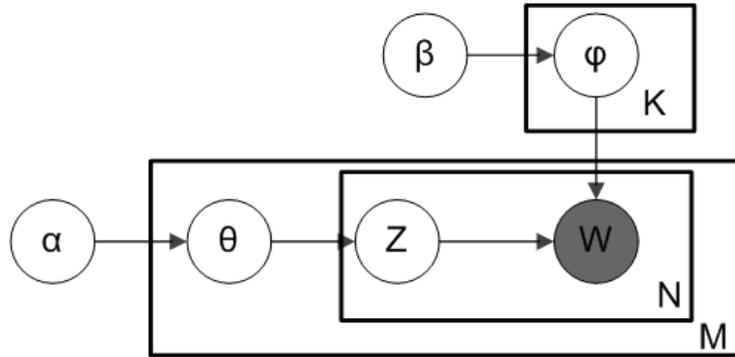
(2) 토픽 모델링(Topic Modeling)

토픽 모델링(Topic Modeling)이란, 방대한 양의 문서로부터 토픽(주제어)을 추출하기 위한 확률 모델의 한 기법으로(Guo, Barnes, & Jia, 2017), 각 문서를 토픽과 단어의 집합으로 가정하고, 해당 문서 내 토픽과 그 토픽을 구성하는 단어의 중요도를 확률적으로 제시하는 방법이다(안병대, 신동원 & 이한준, 2022). 토픽 분석은 문서 집합 내에서 동시 출현 빈도가 높은 단어들을 기준으로 유사한 주제에 따라 문서들을 그룹화한다. 특히, 개별 문서와 주제는 일대일 매칭의 개념이 아닌 여러 주제를 다룰 수 있다는 점을 가정할 때, 소셜 미디어 내 단어들의 집합을 토대로 토픽을 추출할 수 있으며, 문서 집합 내에서 해당 토픽의 출현 빈도는 그것에 대한 관심 정도를 반영한다고 볼 수 있다(김정수, 이석준, 2016).

잠재 디리클레 할당(Latent Dirichlet Allocation, LDA) 기법은 토픽 모델링에서 범용되는 기법으로, 기존 LSI, pLSI 토픽 모델링과는 달리 비구조적인 문서에 대하여 각 문서에 어떤 주제들이 존재하는지를 서술하는 확률에 기반한 토픽 모델링(Topic Modeling) 기법이다. 각 문서는 주요 통계적 관계를 보존하면서, 말뭉치(Corpus)의 효율적인 연산을 위해 구성 요소들에 대한 최적의 주제를 찾는다(Blei, Ng, & Jordan, 2003)

LDA는 디리클레 확률 기반의 토픽 모델링을 통해 문서 내에서 어떤 토픽들이 어떠한 비율로 구성되어 있는지를 나타내어 주며(Blei et al, 2003), 대량의 문서 데이터를 체계적이고 통계적으로 요약함으로써 텍스트 문서 집합을 숫자로 이루어진 토픽과 단어의 벡터 집합으로 변환시켜주기 때문에 최근 온라인 기업리뷰를 분석하는 연구에 많이 활용되고 있다(안병대 등, 2022). LDA 알고리즘의 문서 생성 및 토픽 산출 모델을 도식화하면 <그림 1>과 같다.

<그림 1> LDA Plate Notation



자료 : Latent Dirichlet Allocation (Blei et al. 2003)

M: 문서의 갯수

K: 주제의 갯수

N: 문헌에 속한 단어의 갯수

θ : 문헌의 주제 분포

φ : 주제의 단어 분포

Z: 해당 단어가 속한 주제의 번호

제 2 절 선행연구 분석

(1) ESG 등급과 기업성과에 관한 연구

기존 연구에서는 기업의 비재무적 평가 요소인 ESG 등급을 통해 기업가치를 추산하였다(오상희, 2021; 김윤경, 2020; 임옥빈, 2016). 오상희(2021)는 2011년 ~ 2020년 코스피 기업을 대상으로, 신용 등급과 ESG Total 등급을 통해 기업 가치 관련성을 표준화된 회귀계수 비교를 통해 분석하였다. 임옥빈(2016)은 ESG 점수가 기업가치에 미치는 영향에 대하여 연구하였다. 김윤경(2020)은 Bloomberg ESG Disclosure Score를 활용하여, 2011년 ~ 2017년 유가증권 상장 기업을 대상으로 비재무적 정보 공시의 재무적 성과와 기업가치에 대한 영향을 실증 분석하였다. 선행 연구에서는 비재무적 가치인 ESG 등급을 통해 기업 가치라는 재무적인 가치를 추산하고자 하였으나, ESG 경영을 위한 방안을 제시하지 못하였다는 한계점이 있다.

본 연구에서는 ESG 등급 중에서도 사회(S) 등급에 초점을 맞추어, 사회(S) 수준별로 내부 임직원들의 직무 만족, 불만족 요인을 차별화하여 찾아내고, 이를 인적자원 측면에서 기업의 지속가능한 사회 책임 경영에 대한 시사점을 제시하는 HR Analytics를 수행한다는 점에서 기존 선행 연구들과 차별점을 가진다.

(2) 기업의 ESG활동과 직무만족에 관한 연구

직무만족(Job Satisfaction)이란, 개인이 직무와 관련하여 겪은 다양한 경험에 대한 태도로서, 경험에 대한 만족감 또는 긍정적인 감정상태를 의미한다(Locke, 1976). 기업의 ESG 활동과 직무만족에 대한 연구는 ESG 활동에 대한

근로자의 인식이 직무만족에 어떤 영향을 미치는지를 연구하였다(석정이, 김민정 & 이수열, 2016; 천민호, 설경진, 2022; 한승호, 김문섭 & 서철승 2022; Jin, Kim, 2022; Takimoto & Srisuphaolarn, 2021; Seo, Oh & Kim, 2021).

석정이 등(2016)은 택배 물류 종사자의 CSR(Corporate Social Responsibility, 기업의 사회적 책임) 지각이 조직효과성에 주는 영향을 271명의 택배 물류 종사자를 대상으로 실시한 설문조사 결과를 활용하여 분석하였다. 천민호 외(2022)는 호텔 직원의 ESG 활동 인식이 조직자부심과 조직유효성에 미치는 영향에 대하여 특급호텔 직원 218명에 대해 온라인 설문 조사한 결과를 활용하여 연구하였다. 한승호 등(2022)은 중소기업의 ESG 활동이 근로자들의 자각된 기업이미지와 직무만족에 미치는 영향에 대해 ESG 활동을 실시하고 있는 중소기업의 근로자 383명을 대상으로 온,오프라인 설문조사를 실시한 결과를 연구에 활용하여 연구하였다. 선행 연구들에서는 대상자가 특정된 설문조사를 통해 연구가 이루어졌다는 점에서 수집된 표본의 수가 한정되어 있기 때문에, 일반화된 합의점을 도출해내기에는 다소 제한적이라는 한계점이 있다. 또한, 설문조사는 미리 측정할 변수에 대한 질문이 정해져 있기 때문에 제한된 결과만 수집된다는 한계점이 있다(이중서, 김성근 & 강주영; Coughlan, Cronin, & Ryan 2009에서 재인용).

Jin, Kim(2022)은 혁신적인 조직문화와 자발적 직무설계(Job Crafting)를 매개로 조직 구성원의 ESG 활동 인식이 직무 성과에 미치는 영향을 검증하였다. Takimoto & Srisuphaolarn(2021)은 일본과 태국의 Z세대 대학생을 대상으로 ESG에 대한 인식이 그들의 구직 활동에 어떤 영향을 미치는 지에 대하여 연구하였다. Seo 등(2022)은 직장인 341명의 온라인 설문 데이터를 통해 ESG를 기반으로 직무만족도, 조직신뢰도, 이직의도 간의 관계에 대하여 연구하였다. 선행 연구에서는 기업의 ESG 활동을 평가하는 기준이 연구자의 설계에 따라 각기 다른 설문 수집이 이루어졌다는 점에서 다소 객관성을 가지기가 어렵다. 본 연구에서는 한국ESG기준원(KCGS)에서 실시한, 표준화된 ESG 등급을 활

용하였다는 점과, 기업의 전, 현직자들이 작성한 실제 기업 리뷰를 약 400개 회사의 10만 건에 달하는 빅데이터를 분석하였다는 점에서 표본에 대한 대표성을 향상시킬 수 있다(김동욱, 강주영 & 임재익, 2016).

(3) 온라인 기업 리뷰에 관한 연구

온라인 기업 리뷰란, 자신이 근무한 이력이 있거나 현재 근무 중인 기업에 대한 직원들의 평가를 뜻한다(최진욱, 신동원 & 이한준 2021). 미국의 글래스도어(Glassdoor.com)는 대표적인 기업 리뷰 데이터 공유 플랫폼이다. 글래스도어는 2008년 미국에서 서비스가 시작되었으며, 익명에 기반하여 회사와 상사를 리뷰하고 이를 공유한다.

잡플래닛(Jobplanet.co.kr)은 한국판 글래스도어로, 국내 대표적인 기업 리뷰 데이터 공유 플랫폼이다. 잡플래닛은 2012년 설립된 기업정보 서비스 제공 기업으로, 기업의 전·현직자가 해당 기업에 대해 총 5개의 카테고리(승진 기회 및 가능성, 복지 및 급여, 업무와 삶의 균형, 사내문화, 경영진)에 대해 5점 만점 척도의 정량적 평가를 하고, 4개의 문항(총평(제목), 장점, 단점, 바라는 점)에 대해 리뷰 데이터를 남긴다. 전통적으로, 구직자에게 취업을 결정하는 데 결정적인 요소인 근무 환경이나 사내 문화는 지인 등 기업의 내부 관계자를 통해 구전(Word of Mouth)으로 전해지는 것이 유일한 수단이었으나, 최근에는 잡플래닛이나 블라인드(Blind)와 같은 익명 기반 리뷰 공유 플랫폼을 통해 회사 내 모든 직원경험(Employee Experience)이 공유되고, 해당 기업의 정보에 대해 쉽게 접할 수 있게 되었다. 익명 기반이라는 특성상 악의적인 의도로 남긴 허위 리뷰들이 있을 수 있으며 이에 대한 검증이 부족하고, 개인의 특성이 고려된 다소 주관적인 정보라는 평들이 있으나, 대부분의 사용자들은 온라인 기업 리뷰를 신뢰하고 가치 있게 생각한다(Larkin, 2015). 또한, 인터넷이라는 익

명성이 보장된 공간에서 작성되었기 때문에 연구자의 주관의 개입할 여지가 적어진다는 장점을 가지고 있다(황용석, 2007).

온라인 기업 리뷰는 최근 인사관리 분야의 연구에서도 활발히 사용되고 있다(안병대 등, 2022). 이는 온라인 기업 리뷰 데이터가 기존 설문조사 데이터보다 매우 적은 시간과 비용으로 수집 가능하며, 그 규모가 훨씬 크기 때문이다(최기철, 이상용, 2018).

온라인 기업 리뷰에 대한 분석은 주로 직무만족, 직무불만족 요인에 대해 파악(이중서 등, 2017; 백혜연 & 박용석, 2019; Kim, 2021)하고 이직 선행 요인 파악 등을 통해 직원 이탈을 방지(최진욱 등, 2021; Sull, Sull & Zweig, 2022)하고자 하는 연구들이 진행되어 왔다.

이중서 등(2017)은 우세분석 방법론과 LDA 토픽모델링 기법을 통해 기업 리뷰 데이터를 분석하고, 산업군별 직무만족, 존속요인 및 직무불만족, 이직 요인에 대해 연구하였다. 백혜연 & 박용석(2019)은 기업 리뷰 텍스트를 분석하여 기업에 대한 불만 내용들을 구체적으로 확인하고, 불만 분야에 대한 의미사전목록을 제시하고 부분문법그래프(LGG)를 구축하였다. Kim(2021)은 COVID-19 여파에 따른 직무만족 요인의 패턴 변화를 파악하기 위해, DMR 토픽모델링 방법론을 통해 잡플래닛 기업리뷰 데이터 분석을 실시하였다. 최진욱 등(2021)은 잡플래닛 IT 기업 리뷰 데이터에 토픽모델링 기법을 적용하고, 이를 기반으로 기업의 차년도 이직률을 예측하는 머신러닝 모형을 제안하고, IT 인력의 이직률에 영향력을 갖는 직원 만족 및 불만족 요인을 각각 도출하여 제시하였다. Sull et al. (2022)은 대퇴사시대가 도래함에 따라, 임직원들이 이직하는 사유를 이해하고 직원 이탈을 방지하기 위해 Glassdoor.com의 약 140만 개 이상의 리뷰 데이터를 분석하였다. 선행 연구들에서는 기업 리뷰 데이터를 분석하여 직무만족, 불만족 요인에 대해 파악하였으나, ESG 경영의 측면에서 기업의 사회 책임 경영에 대한 방안을 제시하지 못하였다는 한계점이 있다.

본 연구는 기업 리뷰 데이터와 ESG 경영을 연계하였다는 점에서 의의를 가진다.

온라인 기업 리뷰 데이터를 활용한 국내 연구는 주로 2010년대 후반에 들어서 진행되어 온 것을 알 수 있는데, 그 이유는 각종 텍스트 마이닝 기법을 활용하여 자연어 처리를 할 수 있을 만큼 데이터가 축적된 시점이 비교적 최근이며, 그 배경에는 이직에 대한 관심이 높아지고, 업무 능력과 이직 횟수는 비례한다고 인식되는 대퇴사시대(Great Resignation)가 도래하였기 때문으로 파악할 수 있다. 대학생을 제외한 구직자가 관심있는 기업의 리뷰 데이터를 열람하기 위해서는 유료 상품을 구매하거나, 본인 회사의 리뷰 데이터를 먼저 남겨야 권한이 주어지기 때문에, 기업의 전·현직자들의 리뷰 작성 수가 늘게 된 것으로 파악할 수 있다.

(4) LDA 토픽모델링을 활용한 연구

LDA 토픽모델링을 활용한 기존 연구는 주로 비구조적인 대량의 문서에서 토픽을 도출하고, 토픽을 구성하고 있는 키워드를 함께 제시함으로써 해당 토픽이 전체 문서에서 어떤 의미를 가지는지를 제시한다.

Guo et al. (2016)은 16개국에 위치한 25,670개의 호텔에 대한 266,544건의 호텔 리뷰 데이터를 LDA 토픽모델링을 통해 분석하였다. 조현우(2022)는 Batch 공정 모니터링 연구 논문들의 주요 토픽과 트렌드를 분석하기 위해, 1990년부터 2021년 9월까지 해외 저널에 게시된 총 353편을 선정하여 논문의 제목, 키워드, 초록 데이터를 수집하고 LDA 토픽모델링을 수행하였다. Kwon, Ban, Jun, & Kim (2021)은 항공사를 이용했거나 이용의사가 있는 사람들이 주로 이용하는 Skytrax 게시물 중에서 27개 항공사의 약 14,000개 리뷰를 수집하여 LDA 토픽모델링과 빈도 분석을 수행하였다. Jung & Suh (2018)은 잡플래닛 기업 리뷰

35,063건을 LDA 토픽모델링을 통해 직무만족 요인을 확인하고, 산업, 기업, 그룹, 시간대별 요인을 기반으로 직무만족도 정서와 중요도를 측정하였다. 김동욱 등 (2016)은 잡플래닛 기업 리뷰 데이터를 활용해 LDA 토픽모델링과 동시 출현 단어분석을 수행하여 산업군별 직무만족요인의 차이를 분석하였다. 본 연구는 잡플래닛 기업 리뷰를 기업의 ESG 등급 구간별로 분리하고, 각각의 분리된 데이터를 LDA 토픽모델링을 통해 분석함으로써 도출된 토픽의 차이점에 대해 파악하고, 기업의 사회 책임 경영 측면에서 내부 임직원들의 만족도를 높이기 위한 시사점을 제시한다는 측면에서 의미가 있다

제 III 장 본 론

제 1 절 연구문제 및 설계

(1) 연구 문제

본 연구는 2장에서 설명한 이론적 배경과 선행 연구를 바탕으로, 다음과 같은 연구 문제를 설정하고 연구를 진행하였다.

연구문제 1) 기업의 S(Social) 등급에 따라 잡플래닛 기업 리뷰 '장점'에서 도출되는 토픽과 토픽 단어 구성에 차이가 있는가?

연구문제 2) 기업의 S(Social) 등급에 따라 잡플래닛 기업 리뷰 '단점'에서 도출되는 토픽과 토픽 단어 구성에 차이가 있는가?

연구문제 3) 기업의 S(Social) 등급에 따라 잡플래닛 기업 리뷰 '경영진에 바라는 점'에서 도출되는 토픽과 토픽 단어 구성에 차이가 있는가?

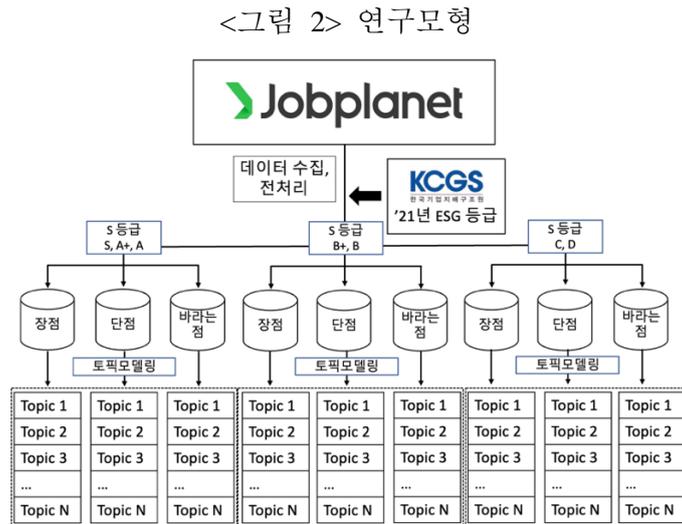
(2) 연구 설계

본 연구는 KCGS에서 평가한 각 기업의 2021년 S(Social) 등급과, 해당 기업의 잡플래닛 리뷰 데이터를 수집하여 LDA 토픽모델링을 수행하고 그 결과를 분석하는 연구로써, 연구 절차는 다음과 같이 진행하였다.

- 1) 데이터 수집(Data Collection)
- 2) 탐색적 데이터 분석(Exploratory Data Analysis)
- 3) 데이터 전처리(Data Preprocessing)
- 4) 데이터 분리(Data Segmentation)
- 5) 토픽 모델링(Topic Modeling)

6) 데이터 시각화(Data Visualization)

연구모형을 도식화하면 아래 <그림 2>와 같다.



제 2 절 연구 대상 및 수집

(1) 연구 대상

본 연구에서는 가장 먼저 2022년 10월 7일 종가 기준 KOSPI 시가 총액 상위 600개 종목을 리스트로 선정하였다. 그 중 우선주와 ETF 종목을 제외한 총 450개 기업을 대상으로 정하고, 최종적으로 지배구조(G)만을 평가하는 금융권 기업과 2022년에 신규 상장하여 2021년 ESG 평가 등급이 존재하지 않는 기업을 제외한 총 404개의 기업을 본 연구의 대상으로 선정하였다.

(2) 연구 수집

1) 데이터 크롤링

본 연구에서는 데이터 수집(크롤링)을 위하여 파이썬 라이브러리인 BeautifulSoup 패키지를 활용하였다. 잡플래닛의 기업 리뷰 데이터는 아래 <그림 3>와 같이 구성되어 있는데, 본 연구에서는 비정형 텍스트 데이터인 '장점', '단점', '경영진에 바라는 점'을 각각 수집하였다. 기업별로 작성된 리뷰의 수가 많은 경우, 최근 작성된 500 개의 리뷰 데이터만을 수집하도록 대상을 선정하였다. 기업 리뷰 데이터가 500 개 이하인 경우에는 작성된 리뷰 전체를 수집하였다.

<그림 3> 잡플래닛 기업 리뷰 데이터 구성



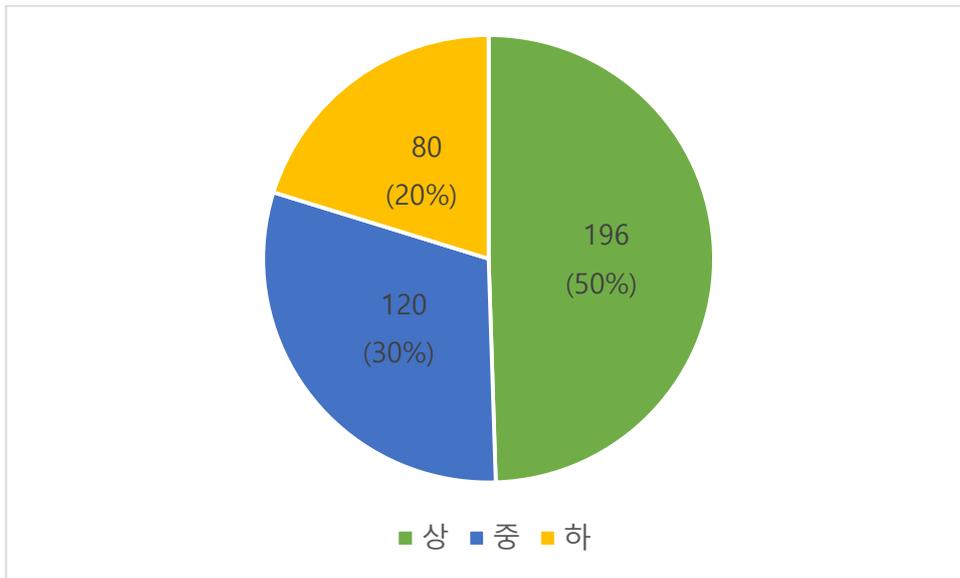
자료 : 잡플래닛(Jobplanet)

또한 그룹 분류를 위해 KCGS 에서 평가한 S(Social) 등급의 S, A+, A 는 '상' 그룹으로 묶고, B+, B 는 '중' 그룹으로, C, D 는 '하' 그룹으로 각각 분류하였다.

2) 탐색적 데이터 분석(Exploratory Data Analysis)

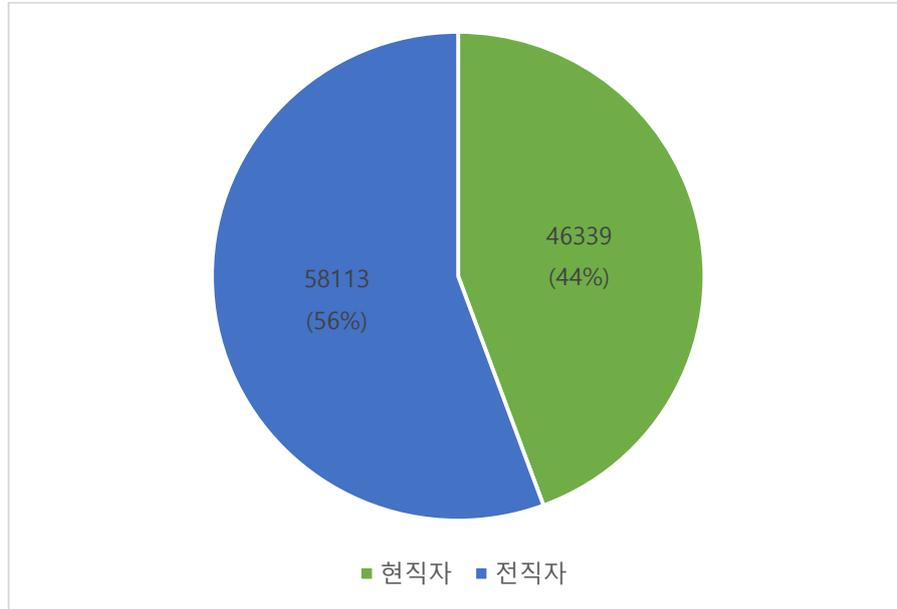
가장 먼저 등급별 기업 분포 <그림 4>를 살펴보면, 총 404 개의 연구 대상 중 '상'에 해당하는 S, A+, A 등급의 기업이 196 개, '중'에 해당하는 B+,B 등급의 기업이 120 개, '하'에 해당하는 등급 C,D 등급의 기업은 80 개였다.

<그림 4> S 등급별 기업 분포



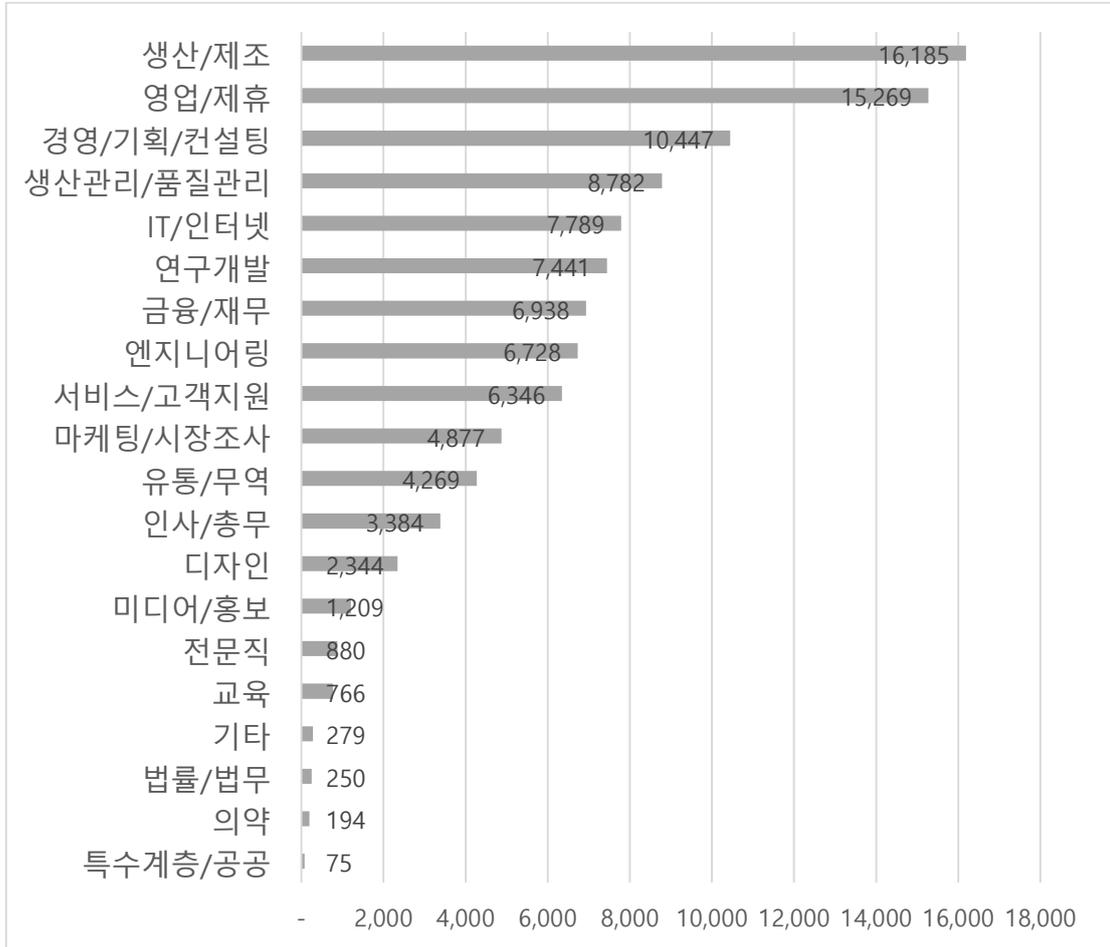
기업 리뷰 데이터의 분포 <그림 5>를 살펴보면, 현직자가 작성한 리뷰는 총 46,339 개이며, 전직자가 작성한 리뷰는 총 58,113 개로 전직자가 작성한 리뷰가 약 11,774 개 더 많은 것을 알 수 있다.

<그림 5> 전/현직자별 잡플래닛 리뷰 갯수



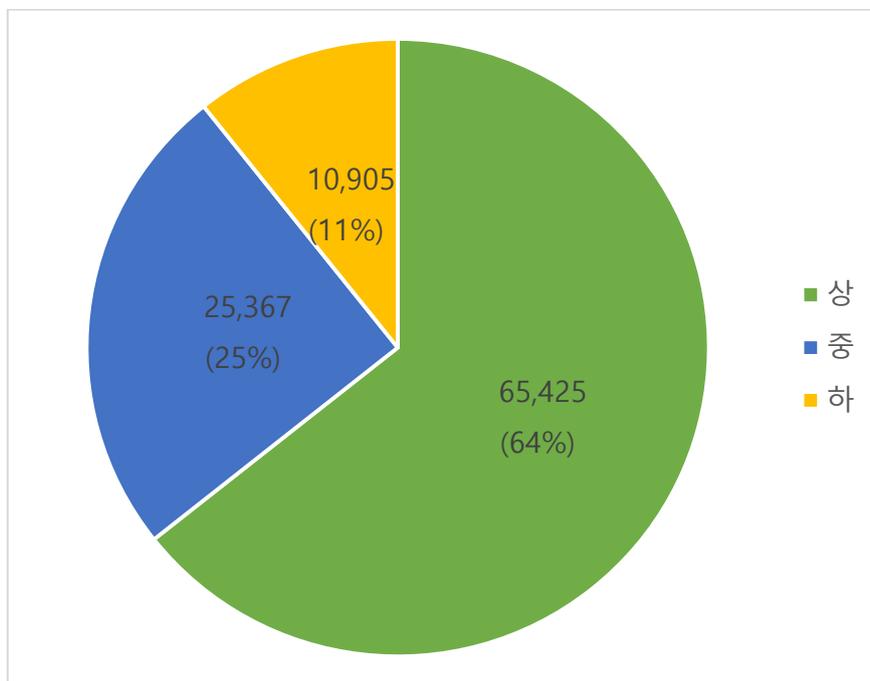
또한 직무별 리뷰 데이터 분포 <그림 6> 를 살펴보면, '생산/제조' 직무의 리뷰가 16,185 개로 가장 많았으며, '영업/채휴'가 15,269 개, '경영/기획/컨설팅' 직무가 10,447 개 순으로 많았다. 리뷰 데이터가 가장 적은 직무는 '특수계층/공공'으로 총 75 개의 리뷰 데이터만이 있었으며, '의약'이 총 194 개, '법률/법무' 직무가 총 250 개 순으로 적었다. 전문직군일수록 리뷰 데이터의 수가 적음을 알 수 있다.

<그림 6> 직무별 잡플래닛 리뷰 갯수



다음으로, 사회(S) 등급별 리뷰 데이터 분포 <그림 7>를 살펴보면, '상'(S, A+, A)에 해당하는 리뷰 데이터가 총 65,425 개로 가장 많았으며, "중"(B+, B)에 해당하는 리뷰 데이터는 총 25,367 개, '하'(C, D)에 해당하는 리뷰 데이터는 총 10,905 개로 가장 적었다.

<그림 7> S 등급별 잡플래닛 리뷰 갯수



제 3 절 연구 방법

(1) 데이터 전처리(Data Preprocessing)

본 연구에서는 LDA 토픽모델링을 수행하기 위해 다음과 같은 데이터 전처리 과정을 거쳤다. 가장 먼저 불용어를 제거하였으며, 한글 외 텍스트는 제거하는 작업을 거쳤다. 이후 파이썬의 KoNLPy 라이브러리에 내장되어 있는 Mecab-ko 형태소 분석기를 활용하여 토큰화를 진행하고, 명사만을 추출하였다.

1) 불용어 제거

불용어란, 매우 빈번하게 출현하지만 텍스트의 의미 구성에는 영향을 미치지 않는 불필요한 어휘들로, 이런 어휘들을 제거해야 비로소 텍스트의 의미를 구성하는 중요한 어휘들을 유효하게 산출할 수 있다(길호현, 2018). 따라서 본 연구에서는 불용어 제거 작업을 거쳐 보다 더 의미있는 토픽을 추출해내고자 하였다. 불용어의 목록은 "텍스트 마이닝을 위한 한국어 불용어 목록 연구(길호현, 2018)" 연구에서 제시한 '의미 구조 분석을 위한 표준 불용어 목록 시안'에 포함된 총 293 개의 단어를 채택하였다. 단어 목록은 아래 <표 4>와 같다.

<표 4> 불용어 목록

가랑, 가지, 각, 간, 갓은, 개, 개국, 개년, 개소, 개월, 개, 거, 거기, 거리, 건, 것, 겨를, 격, 겸, 고, 군, 군데, 권, 그, 그거, 그것, 그곳, 그까짓, 그네, 그녀, 그놈, 그대, 그래, 그래도, 그서, 그러나, 그러니, 그러니까, 그러다가, 그러면, 그러면서, 그러므로, 그러자, 그런, 그런데, 그럼, 그렇지만, 그루, 그리고, 그리하여, 그분, 그이, 그쪽, 근, 근데, 글썸, 글썸요, 기, 김, 나, 나름, 나위, 남짓, 내, 냥, 너, 너희, 네, 네놈, 녀석, 년, 년대, 년도, 놈, 누구, 니, 다른, 다만, 단, 달, 달리, 당신, 대, 대로, 더구나, 더욱이, 데, 도, 동, 되, 두, 두세, 두어, 둥, 듯, 듯이, 등, 등등, 등지, 따라서, 따름, 따위, 판, 때문, 또, 또는, 또한, 리, 마당, 마련, 마리, 만, 만큼, 말, 매, 맨, 명, 몇, 몇몇, 모, 모금, 모든, 무렵, 무슨, 무엇, 뭐, 뭇, 미터, 밋, 바, 바람, 바퀴, 박, 발, 발짝, 번, 별, 법, 별, 본, 부, 분, 뽀, 뽀, 살, 새, 서너, 석, 설, 섬, 세, 세기, 썸, 씌네, 수, 순, 스무, 승, 시, 시간, 식, 씨, 아, 아냐, 아니, 아니야, 아무, 아무개, 아무런, 아아, 아이, 아이고, 아이구, 야, 약, 양, 애, 어, 어느, 어디, 어머니, 언제, 에이, 엔, 여기, 여느, 여러, 여러분, 여보, 여보세요, 여지, 역시, 예, 옛, 오, 오랜, 오히려, 온, 온갖, 올, 왜냐하면, 웬, 외, 요, 우리, 원, 월, 웬, 위, 음, 응, 이, 이거, 이것, 이곳, 이놈, 이래, 이런, 이런저런, 이른바, 이리하여, 이쪽, 일, 일대, 임마, 자, 자기, 자네, 장, 저, 저것, 저기, 저놈, 저런, 저쪽, 저편, 저희, 적, 전, 점, 제, 조, 주, 주년, 주일, 줄, 중, 즈음, 즉, 지, 지경, 지난, 집, 짝, 쪽, 쯤, 차, 참, 채, 척, 찻, 체, 초, 총, 측, 치, 큰, 킬로미터, 타, 터, 턱, 툄, 통, 투, 판, 퍼센트, 편, 평, 폰, 하기야, 하긴, 하물며, 하지만, 한, 한두, 한편, 허허, 현, 현, 호, 혹은, 회

자료 : 텍스트 마이닝을 위한 한국어 불용어 목록 연구 (길호현, 2018)

2) 토큰화(Tokenization)

텍스트에서 토큰(Token)이란, 상황에 따라 다르지만 보통 '의미를 가지는 최소 단위'로 정의한다. 주어진 코퍼스(Corpus)를 토큰 단위로 나누는 작업을 토큰화(Tokenization)라고 한다(Yoo, 2021). 한국어는 교착어라는 특성을 가지고 있기 때문에, 의미를 가진 단어에 어미, 조사 등의 불용어가 붙게 된다. 따라서 한국어를 토큰화 할 때는 주로 형태소 단위로 토큰화를 실시하게 된다. 파이썬 라이브러리 KoNLPy 는 2014 년 서울대학교에서 개발한 한국어 정보처리 도구이다(박은정, 조성준, 2014). KoNLPy 에는 Kkma, Hannanum, Mecab-ko 등 여러가지 형태소 분석기가 내장되어 있는데, 본 연구에서는 한국어 형태소 분석기 중 가장 성능이 뛰어나다고 알려져 있는 Mecab-ko 형태소 분석기를 사용하여 토큰화를 진행하였다. Mecab-ko 분석기는 일본어 형태소 분석기 Mecab 을 한국어용으로 포팅한 버전으로, 실행 속도가 가장 빠르고 분석 품질도 양호하다(강형석, 양장훈, 2018).

앞선 전처리 과정과 토큰화 과정을 거친 텍스트 데이터에서, 본 연구에서는 명사 데이터만 추출하여 활용하였다. 동사, 형용사는 어미에 따라 그 형태가 매우 다양해지기 때문에 토크모델링을 활용하기 적합하지 않다고 판단하였기 때문이다.

(2) LDA 토픽 모델링(Topic Modeling)

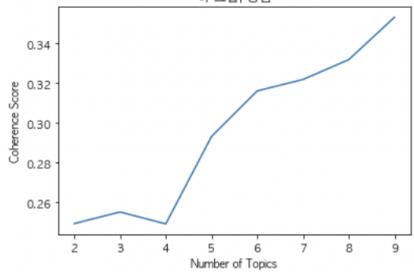
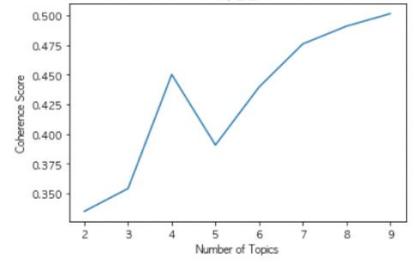
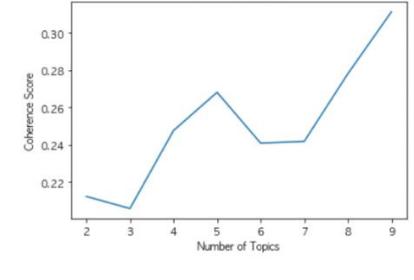
LDA 토픽 모델링을 시행하기에 앞서, 토픽의 갯수(Number of Topic)는 연구자가 직접 설정해야 하는 하이퍼파라미터(Hyperparameter)이다. 최적의 토픽 수 선정을 위한 지표로써 일반적으로 혼란도(Perplexity)와 일관성(Coherence) 지표를 사용하는데(Chang, Gerrish, Wang, Boyd-Graber & Blei, 2009) 본 연구에서는 최적의 토픽 수 선택을 위해 일관성 지표(Coherence Score)를 산정하였다. 일관성 지표란, 토픽이 의미론적으로 얼마나 일관성이 있는지를 판단하는 지표로서, 그 값이 높을수록 의미론적으로 일관성이 있음을 의미한다(Newman, Lau, Grieser & Baldwin, 2010). 상, 중, 하 각 등급별 장점, 단점, 경영진에 바라는 점 총 9 개의 일관성 지표를 산출하였으며, 토픽 갯수의 범위는 최소 2 개, 최대 10 개로 설정하였다. 각 리뷰별 최적의 토픽 갯수 및 일관성 지표는 아래 <표 5>와 같다.

<표 5> 각 리뷰별 최적의 토픽 갯수 및 일관성 지표

Rank	Type	Number of Topic	Coherence Score																		
전체	장점	8	<table border="1"> <caption>전체, 장점 - 일관성 지표</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.18</td></tr> <tr><td>3</td><td>0.27</td></tr> <tr><td>4</td><td>0.26</td></tr> <tr><td>5</td><td>0.31</td></tr> <tr><td>6</td><td>0.32</td></tr> <tr><td>7</td><td>0.34</td></tr> <tr><td>8</td><td>0.37</td></tr> <tr><td>9</td><td>0.35</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.18	3	0.27	4	0.26	5	0.31	6	0.32	7	0.34	8	0.37	9	0.35
Number of Topics	Coherence Score																				
2	0.18																				
3	0.27																				
4	0.26																				
5	0.31																				
6	0.32																				
7	0.34																				
8	0.37																				
9	0.35																				

전체	단점	9	<p>전체, 단점</p> <table border="1"> <caption>Data for '전체, 단점' graph</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.38</td></tr> <tr><td>3</td><td>0.41</td></tr> <tr><td>4</td><td>0.43</td></tr> <tr><td>5</td><td>0.46</td></tr> <tr><td>6</td><td>0.47</td></tr> <tr><td>7</td><td>0.50</td></tr> <tr><td>8</td><td>0.49</td></tr> <tr><td>9</td><td>0.53</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.38	3	0.41	4	0.43	5	0.46	6	0.47	7	0.50	8	0.49	9	0.53
Number of Topics	Coherence Score																				
2	0.38																				
3	0.41																				
4	0.43																				
5	0.46																				
6	0.47																				
7	0.50																				
8	0.49																				
9	0.53																				
전체	경영진에 바라는 점	8	<p>전체, 바라는점</p> <table border="1"> <caption>Data for '전체, 바라는점' graph</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.20</td></tr> <tr><td>3</td><td>0.24</td></tr> <tr><td>4</td><td>0.35</td></tr> <tr><td>5</td><td>0.45</td></tr> <tr><td>6</td><td>0.40</td></tr> <tr><td>7</td><td>0.36</td></tr> <tr><td>8</td><td>0.45</td></tr> <tr><td>9</td><td>0.42</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.20	3	0.24	4	0.35	5	0.45	6	0.40	7	0.36	8	0.45	9	0.42
Number of Topics	Coherence Score																				
2	0.20																				
3	0.24																				
4	0.35																				
5	0.45																				
6	0.40																				
7	0.36																				
8	0.45																				
9	0.42																				
상	장점	8	<p>상 그룹, 장점</p> <table border="1"> <caption>Data for '상 그룹, 장점' graph</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.23</td></tr> <tr><td>3</td><td>0.27</td></tr> <tr><td>4</td><td>0.30</td></tr> <tr><td>5</td><td>0.29</td></tr> <tr><td>6</td><td>0.31</td></tr> <tr><td>7</td><td>0.31</td></tr> <tr><td>8</td><td>0.33</td></tr> <tr><td>9</td><td>0.32</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.23	3	0.27	4	0.30	5	0.29	6	0.31	7	0.31	8	0.33	9	0.32
Number of Topics	Coherence Score																				
2	0.23																				
3	0.27																				
4	0.30																				
5	0.29																				
6	0.31																				
7	0.31																				
8	0.33																				
9	0.32																				
상	단점	8	<p>상 그룹, 단점</p> <table border="1"> <caption>Data for '상 그룹, 단점' graph</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.37</td></tr> <tr><td>3</td><td>0.35</td></tr> <tr><td>4</td><td>0.40</td></tr> <tr><td>5</td><td>0.41</td></tr> <tr><td>6</td><td>0.41</td></tr> <tr><td>7</td><td>0.43</td></tr> <tr><td>8</td><td>0.47</td></tr> <tr><td>9</td><td>0.43</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.37	3	0.35	4	0.40	5	0.41	6	0.41	7	0.43	8	0.47	9	0.43
Number of Topics	Coherence Score																				
2	0.37																				
3	0.35																				
4	0.40																				
5	0.41																				
6	0.41																				
7	0.43																				
8	0.47																				
9	0.43																				

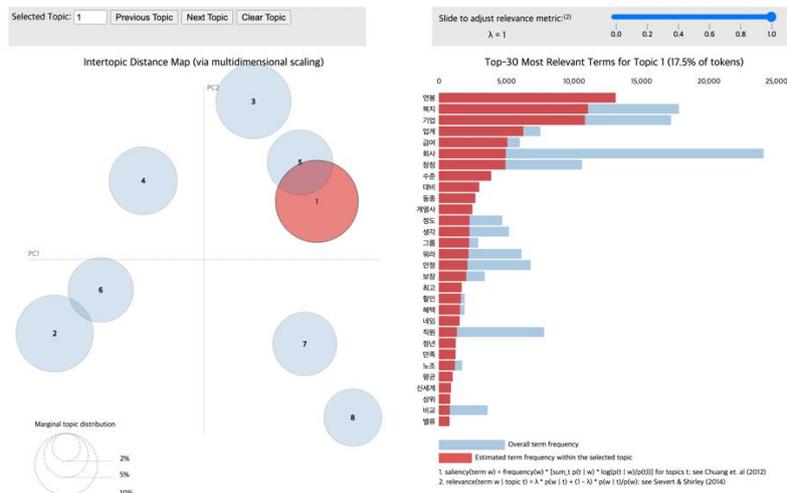
상	경영진에 바라는 점	8	<p>상 그룹, 바라는점</p> <table border="1"> <caption>상 그룹, 바라는점</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.20</td></tr> <tr><td>3</td><td>0.28</td></tr> <tr><td>4</td><td>0.35</td></tr> <tr><td>5</td><td>0.32</td></tr> <tr><td>6</td><td>0.38</td></tr> <tr><td>7</td><td>0.40</td></tr> <tr><td>8</td><td>0.42</td></tr> <tr><td>9</td><td>0.40</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.20	3	0.28	4	0.35	5	0.32	6	0.38	7	0.40	8	0.42	9	0.40
Number of Topics	Coherence Score																				
2	0.20																				
3	0.28																				
4	0.35																				
5	0.32																				
6	0.38																				
7	0.40																				
8	0.42																				
9	0.40																				
중	장점	7	<p>중 그룹, 장점</p> <table border="1"> <caption>중 그룹, 장점</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.27</td></tr> <tr><td>3</td><td>0.27</td></tr> <tr><td>4</td><td>0.31</td></tr> <tr><td>5</td><td>0.31</td></tr> <tr><td>6</td><td>0.37</td></tr> <tr><td>7</td><td>0.37</td></tr> <tr><td>8</td><td>0.35</td></tr> <tr><td>9</td><td>0.36</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.27	3	0.27	4	0.31	5	0.31	6	0.37	7	0.37	8	0.35	9	0.36
Number of Topics	Coherence Score																				
2	0.27																				
3	0.27																				
4	0.31																				
5	0.31																				
6	0.37																				
7	0.37																				
8	0.35																				
9	0.36																				
중	단점	9	<p>중, 단점</p> <table border="1"> <caption>중, 단점</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.34</td></tr> <tr><td>3</td><td>0.44</td></tr> <tr><td>4</td><td>0.45</td></tr> <tr><td>5</td><td>0.49</td></tr> <tr><td>6</td><td>0.50</td></tr> <tr><td>7</td><td>0.49</td></tr> <tr><td>8</td><td>0.50</td></tr> <tr><td>9</td><td>0.50</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.34	3	0.44	4	0.45	5	0.49	6	0.50	7	0.49	8	0.50	9	0.50
Number of Topics	Coherence Score																				
2	0.34																				
3	0.44																				
4	0.45																				
5	0.49																				
6	0.50																				
7	0.49																				
8	0.50																				
9	0.50																				
중	경영진에 바라는 점	9	<p>중 그룹, 바라는점</p> <table border="1"> <caption>중 그룹, 바라는점</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.22</td></tr> <tr><td>3</td><td>0.16</td></tr> <tr><td>4</td><td>0.34</td></tr> <tr><td>5</td><td>0.34</td></tr> <tr><td>6</td><td>0.29</td></tr> <tr><td>7</td><td>0.28</td></tr> <tr><td>8</td><td>0.32</td></tr> <tr><td>9</td><td>0.35</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.22	3	0.16	4	0.34	5	0.34	6	0.29	7	0.28	8	0.32	9	0.35
Number of Topics	Coherence Score																				
2	0.22																				
3	0.16																				
4	0.34																				
5	0.34																				
6	0.29																				
7	0.28																				
8	0.32																				
9	0.35																				

하	장점	9	<p>하 그룹, 장점</p>  <table border="1"> <caption>Data for 하 그룹, 장점</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.25</td></tr> <tr><td>3</td><td>0.26</td></tr> <tr><td>4</td><td>0.25</td></tr> <tr><td>5</td><td>0.29</td></tr> <tr><td>6</td><td>0.32</td></tr> <tr><td>7</td><td>0.32</td></tr> <tr><td>8</td><td>0.33</td></tr> <tr><td>9</td><td>0.35</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.25	3	0.26	4	0.25	5	0.29	6	0.32	7	0.32	8	0.33	9	0.35
Number of Topics	Coherence Score																				
2	0.25																				
3	0.26																				
4	0.25																				
5	0.29																				
6	0.32																				
7	0.32																				
8	0.33																				
9	0.35																				
하	단점	9	<p>하, 단점</p>  <table border="1"> <caption>Data for 하, 단점</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.34</td></tr> <tr><td>3</td><td>0.36</td></tr> <tr><td>4</td><td>0.45</td></tr> <tr><td>5</td><td>0.39</td></tr> <tr><td>6</td><td>0.44</td></tr> <tr><td>7</td><td>0.47</td></tr> <tr><td>8</td><td>0.48</td></tr> <tr><td>9</td><td>0.50</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.34	3	0.36	4	0.45	5	0.39	6	0.44	7	0.47	8	0.48	9	0.50
Number of Topics	Coherence Score																				
2	0.34																				
3	0.36																				
4	0.45																				
5	0.39																				
6	0.44																				
7	0.47																				
8	0.48																				
9	0.50																				
하	경영진에 바라는 점	9	<p>하 그룹, 바라는점</p>  <table border="1"> <caption>Data for 하 그룹, 바라는점</caption> <thead> <tr> <th>Number of Topics</th> <th>Coherence Score</th> </tr> </thead> <tbody> <tr><td>2</td><td>0.21</td></tr> <tr><td>3</td><td>0.21</td></tr> <tr><td>4</td><td>0.25</td></tr> <tr><td>5</td><td>0.27</td></tr> <tr><td>6</td><td>0.24</td></tr> <tr><td>7</td><td>0.24</td></tr> <tr><td>8</td><td>0.28</td></tr> <tr><td>9</td><td>0.31</td></tr> </tbody> </table>	Number of Topics	Coherence Score	2	0.21	3	0.21	4	0.25	5	0.27	6	0.24	7	0.24	8	0.28	9	0.31
Number of Topics	Coherence Score																				
2	0.21																				
3	0.21																				
4	0.25																				
5	0.27																				
6	0.24																				
7	0.24																				
8	0.28																				
9	0.31																				

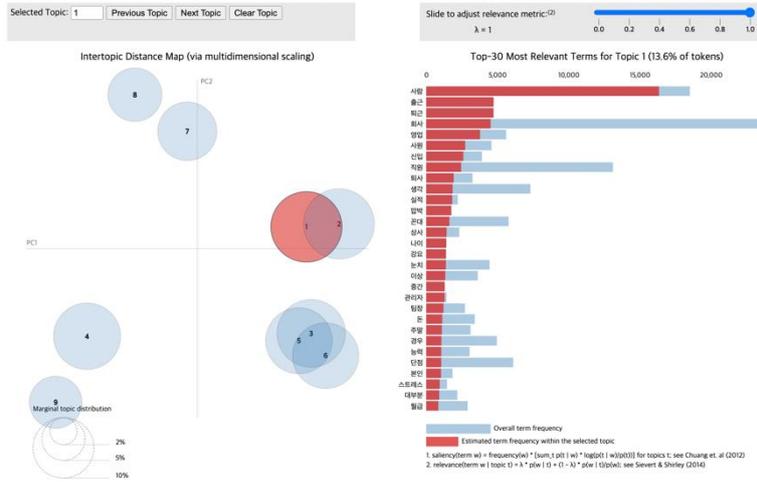
(3) LDAvis 를 활용한 LDA 토픽 모델링 시각화

LDAvis 는 LDA 토픽모델링의 결과를 차원 축소(PCA, Principal Component Analysis)와 키워드 추출 방법을 통하여, 토픽을 구성하는 키워드 및 토픽 간의 관계를 한 눈에 파악할 수 있도록 시각화해주는 도구이다(Sievert & Shirley, 2014). 화면 좌측의 2 차원 평면에서 각 원은 개별 토픽을 의미하며, 원의 크기는 전체 문서에서 해당 토픽이 차지하는 비중을 의미한다. 원이 커지면 커질 수록 전체 문서에서 해당 토픽이 차지하는 비중이 크다는 것을 의미한다. 우측 가로 막대 그래프는 해당 토픽 내에서 차지하는 단어들의 비중을 의미한다. 막대 그래프의 길이가 길수록, 해당 단어가 토픽 내에서 언급되는 빈도수가 높음을 의미한다. 본 시각화 도구는 HTML 상에서 사용자의 조작에 따라 화면이 상호작용하는 시각화 도구로서, 지면의 한계 상 모든 상호작용 내용을 담을 수는 없기 때문에 대표적인 화면만 나타내도록 하였다. 아래 <그림 8> ~ <그림 10>은 전체 리뷰 데이터 장점, 단점, 경영진에 바라는 점에 대해 LDA 토픽모델링을 시행하고 시각화를 한 결과이다.

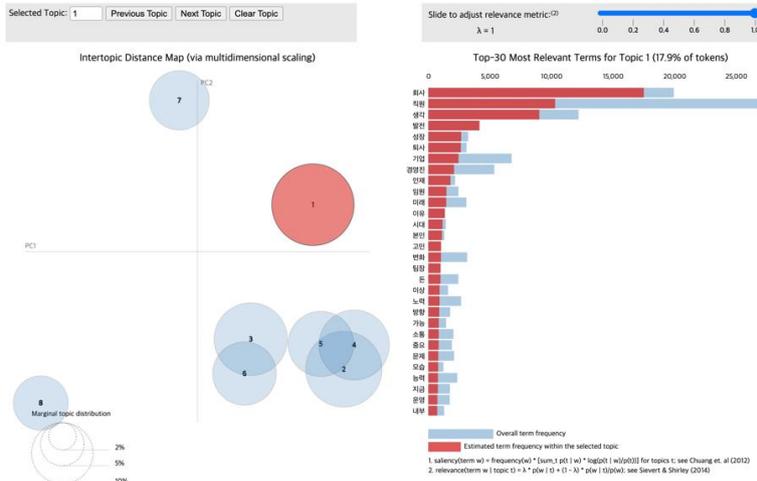
<그림 8> 장점(전체) LDA 토픽모델링 시각화



<그림 9> 단점(전체) LDA 토픽모델링 시각화



<그림 10> 경영진에 바라는 점(전체) LDA 토픽모델링 시각화



자료 : LDAvis (Sievert & Shirley, 2014)

제 4 절 토픽모델링 결과

(1) 기업 리뷰 '장점' 등급별 토픽 모델링 결과

'장점' 리뷰에 대해 전체 데이터 및 상, 중, 하 등급별로 구분하여 LDA 토픽모델링을 실시한 결과는 아래 <표 6> ~ <표 9>와 같다.

전체 리뷰에 대한 LDA 토픽모델링 결과, '동종업계 대비 우위', '자유로운 휴가 사용', '경영 환경', '조직문화', '영업 환경', '업무 강도', '근무 환경', '급여 및 보상'이 주요 토픽으로 언급되었다.

등급별 특징을 살펴보면, 일과 삶의 균형(Work & Life Balance)을 뜻하는 '워라벨'이라는 키워드는 '상', '중' 등급에서만 언급되는 것을 확인할 수 있었다. 이는 '하' 등급에서는 토픽 내 키워드로 언급되지 않았음을 동시에 확인하였는데, 이러한 점으로 미루어 보면 '하' 등급에 속하는 기업에서는 가족친화적 제도들이 수립되고, 실질적으로 시행이 되고 있는지 점검을 통해 일과 삶의 균형을 이루는 가족친화경영이 필요할 것으로 보인다.

마찬가지로 '상', '중' 등급에서는 수평적인 조직문화가 주요 토픽으로 '수평', '자유', '분위기', '문화' 등의 키워드가 함께 언급되나, '하' 등급에서는 수평적인 조직문화에 대한 토픽 언급이 없는 점을 확인할 수 있었다. '하' 등급에 속하는 회사에서는 수평적인 조직문화를 갖추기 위해 구성원의 인식과 행동 개선을 유도하기 위한 조직문화 활동 등이 필요할 것으로 보인다.

마지막으로, '하' 등급에서만 나타나는 장점의 특징은 비교적 구체적인 키워드를 포함한다는 것이다. '회사 소재지'가 장점의 주요 토픽으로 언급되면서, '위치', '서울' 등의 키워드가 함께 언급되는 것을 확인할 수 있었는데, 이는 '하' 등급에 속하는 기업들이 주로 제조업을 영위하며, 리뷰에서 가장 많이 차지하고 있는 직무가 '생산/제조'인 점을 반영하면

제조업 특성상 사업장이 주로 지방에 위치하고 있기 때문에, 임직원들이 수도권 근무를 선호한다고 이해할 수 있다. 비슷한 맥락에서 '아침', '점심', '저녁', '식사', '기숙사', '통근버스' 등의 키워드가 장점으로 동시에 언급되는 것 또한 사업, 직무 특성이 반영된 것으로 이해할 수 있다. 따라서 기업에서는 신체적 안전과 편의, 정서적 안정, 심리적 안정에 최적화된 근무 환경을 조성하여 제품과 서비스의 질 향상, 생산성 증대와 근로자의 사기 진작을 도모하는 안전보건 거버넌스가 구축되어야 할 것이다(한국 ESG 기준원, 2021).

<표 6> 전체 그룹 '장점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
1	동종업계 대비 우위	17.5%	연봉 8.9%	복지 7.5%	기업 7.4%	업계 4.3%	급여 3.5%	회사 3.4%	장점 3.4%	수준 2.6%	대비 2.0%	동종 1.8%
2	자유로운 휴가 사용	15.3%	연차 13.4%	자유 8.3	사용 7.5%	가능 6.8%	눈치 6.3%	퇴근 4.0%	휴가 3.8%	야근 3.0%	출퇴근 2.8%	근무 2.1%
3	경영환경	14.4%	회사 8.7%	안정 3.9%	기업 3.3%	사업 2.2%	가능 2.1%	성장 2.1%	교육 2.0%	기회 1.8%	경험 1.6%	규모 1.6%
4	조직문화	11.9%	분위기 11.2%	문화 6.6%	부서 3.4%	사람 3.2%	직원 3.1%	회사 2.5%	기업 2.4%	수평 2.2%	조직 1.9%	업무 1.9%
5	영업환경	11%	사람 10.1%	회사 4.4%	영업 2.8%	사원 2.1%	중견 2.0%	매출 1.9%	신입 1.8%	장점 1.6%	이상 1.4%	압박 1.3%
6	업무환경	10.9%	업무 13.9%	근무 7.5%	강도 4.7%	환경 3.5%	초봉 3.1%	위치 2.7%	서울 2.3%	본사 1.9%	가능 1.8%	장점 1.6% [^]
7	근무조건	10.5%	제공 5.9%	점심 3.3%	식당 2.6%	복지 2.5%	회사 2.3%	지원 2.2%	기숙사 2.1%	버스 1.9%	저녁 1.8%	사내 1.8%
8	보상체계	8.5%	월급 8.4%	돈 6.5%	지원 5.5%	수당 3.3%	복지 2.9%	지급 2.8%	장점 2.5%	윗사람 1.9%	상여금 1.7%	명절 1.4%

<표 7> 상 그룹 '장점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
1	경영환경	15.7%	안정 3.9%	회사 3.9%	업무 2.7%	영업 2.4%	기업 2.3%	가능 2.2%	경험 2.2%	사업 2.1%	성장 2.1%	교육 2.0%
2	자유로운 휴가사용	15.3%	연차 13.6%	자유 9.0%	사용 7.8%	눈치 6.6%	가능 6.0%	휴가 5.8%	분위기 2.2%	퇴근 2.2%	복지 2.1%	출퇴근 1.9%
3	동종업계 대비 우위	15.1%	기업 11.0%	복지 8.4%	연봉 6.8%	업계 5.7%	회사 3.5%	수준 3.0%	급여 2.7%	대비 2.4%	계열사 2.1%	동종 2.1%
4	조직문화	14.6%	분위기 8.6%	문화 6.5%	사람 6.2%	업무 3.8%	회사 2.7%	직원 2.5%	수평 2.4%	기업 1.8%	조직 1.7%	자유 1.5%
5	일과 삶의 균형	10.7%	근무 9.9%	업무 5.5%	장점 4.9%	환경 4.0%	강도 3.8%	위라벨 3.5%	부서 3.0%	위치 2.5%	본사 2.2%	보장 2.0%
6	근무조건	9.7%	제공 6.0%	돈 5.5%	점심 3.5%	식당 2.6%	사내 2.5%	회사 2.1%	구내 식당 2.0%	건물 1.9%	식사 1.8%	가능 1.7%
7	복리후생	9.6%	복지 11.4%	지원 6.9%	직원 3.2%	할인 3.0%	혜택 2.5%	지급 1.8%	기본 1.8%	포인트 1.4%	제도 1.3%	상여금 1.3%
8	보상체계	9.3%	회사 6.0%	월급 3.6%	연봉 3.0%	생각 2.7%	그룹 2.5%	백화점 2.3%	사람 2.3%	복리 2.3%	이상 2.3%	경험 2.2%

<표 8> 중 그룹 '장점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
1	업무경험	19.8%	업무 8.6%	회사 2.8%	강도 2.4%	가능 2.0%	영업 2.0%	분위기 1.8%	경험 1.8%	사람 1.6%	기회 1.5%	체계 1.4%
2	자유로운 휴가사용	17%	연차 11.4%	자유 8.1%	사용 6.2%	눈치 5.4%	가능 5.2%	퇴근 3.3%	휴가 3.0%	분위기 3.0%	부서 2.4%	위라벨 2.1%
3	동종업계 대비 우위	16.5%	회사 7.9%	기업 6.9%	업계 4.8%	연봉 4.3%	안정 3.6%	수준 1.7%	대비 1.6%	동종 1.5%	월급 1.4%	사업 1.3%
4	근무조건	12.3%	장점 10.0%	돈 3.6%	월급 3.5%	제공 3.4%	위치 3.0%	회사 2.3%	점심 2.1%	서울 2.0%	근무 2.0%	야근 1.9%
5	복리후생	12%	복지 14.3%	지원 4.8%	급여 3.4%	연봉 2.8%	기업 2.6%	직원 1.9%	안정 1.9%	지급 1.7%	수준 1.6%	혜택 1.4%
6	조직문화	11.9%	분위기 4.4%	근무 4.0%	문화 3.8%	사람 2.9%	사내 2.7%	직원 2.5%	식당 2.1%	환경 2.0%	수평 1.8%	건물 1.6%
7	보상체계	11%	생각 5.7%	회사 3.6%	초봉 3.4%	기업 2.5%	연봉 2.4%	사람 2.2%	네임 1.8%	신입 1.5%	사원 1.5%	정규직 1.1%

<표 9> 하 그룹 '장점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
1	자유로운 휴가 사용	19.8%	연차 8.4%	자유 6.2%	분위기 5.0%	사용 4.5%	가능 4.4%	눈치 4.0%	부서 3.3%	업무 3.0%	퇴근 3.0%	사람 2.6%
2	경영환경	15.7%	회사 10.2%	안정 3.9%	기업 2.6%	영업 2.5%	사업 1.7%	매출 1.4%	국내 1.4%	구조 1.3%	채무 1.0%	성장 1.0%
3	업무환경	11.4%	업무 5.9%	기업 2.7%	근무 2.6%	개발 2.1%	회사 2.0%	노력 1.9%	직원 1.9%	복지 1.8%	가능 1.7%	환경 1.5%
4	동종업계 대비 우위	10.5%	연봉 12.7%	업계 7.5%	기업 3.4%	대비 3.0%	초봉 2.9%	동종 2.7%	수준 2.5%	신입 2.5%	복지 2.3%	급여 2.0%
5	근무조건	9.8%	제공 8.5%	지원 5.6%	점심 4.9%	기숙사 3.6%	저녁 3.3%	식사 2.5%	아침 2.5%	버스 2.4%	수당 2.3%	통근 2.2%
6	근무조건2	9%	사람 9.0%	위치 5.9%	회사 4.2%	서울 3.7%	돈 3.7%	식당 2.7%	본사 2.1%	생활 2.0%	직장 1.8%	근무 1.5%
7	복리후생	8.5%	장점 15.4%	기업 4.7%	회사 4.0%	생각 3.8%	정도 2.6%	돈 2.4%	복지 2.3%	업무 2.2%	월급 1.8%	근무 1.7%
8	업무경험	8.3%	직원 4.4%	경험 4.0%	해외 2.8%	현장 2.2%	기회 2.0%	사람 1.9%	업무 1.8%	만족 1.5%	다양 1.4%	복지 1.4%
9	보상체계	7%	월급 4.9%	회사 3.7%	가능 2.1%	직원 1.8%	임금 1.3%	제때 1.3%	건설사 1.1%	대리 1.1%	안정 1.1%	보너스 1.1%

(2) 기업 리뷰 '단점' 등급별 토픽 모델링 결과

'단점' 리뷰에 대해 전체 데이터 및 상, 중, 하 등급별로 구분하여 LDA 토픽모델링을 실시한 결과는 아래 <표 10> ~ <표 13>과 같다.

전체 리뷰에 대한 LDA 토픽 모델링 결과, '근무환경', '급여 및 보상', '성장 전략', '업무 강도', '조직문화', '평가체계', '휴가', '직급체계', '문서업무' 등이 주요 단점 토픽으로 도출되었음을 확인할 수 있다.

등급별로 나타나는 특징을 살펴보면, 가장 먼저 '상', '중' 등급에서만 나타나는 특징으로 '문서 업무'라는 토픽이 추출되었다는 점이다. 문서 업무(Paperwork) 토픽은 주로 '보고', '회의', '단순', '문서', '효율' 등의 키워드들이 함께 언급되고 있는 것을 확인할 수 있는데, 이는 '단순 문서 작업', '보고를 위한 자료 작성', '형식적인 회의'가 많다는 점을 시사한다. 이러한 점에 착안하여 '상', '중'에 해당하는 기업에서는 적절한 권한 위임을 통해 의사결정의 단계를 간소화하고, 각종 문서 업무에 대해 페이퍼리스(Paperless) 캠페인 등의 시행과 더불어, 업무 효율성을 높이고 내부 임직원들의 근무 만족도를 높일 수 있는 스마트 워크 센터 구축, 재택근무 확대 등의 스마트 워크(Smart Work) 환경을 구축해 나가야 할 것으로 보인다(이재성 & 김홍식, 2010).

또한 '하' 등급에서만 나타나는 주요 특징은 임원의 의사결정이 주요 토픽으로 언급된다는 것이다. '임원', '결정', '보수', '능력', '생각' 등의 키워드가 함께 언급되는 것을 확인할 수 있었는데, 업무에 주요 의사 결정권자인 임원의 보수적인 태도에서 비롯한 명확하지 않은 업무지시 및 의사결정이 주요 단점으로 언급된 것을 확인할 수 있다. '하' 등급에 해당하는 기업에서는 임원들의 명확한 의사결정을 통해, 임직원들이 상급자를 신뢰하고 근무할 수 있는 환경을 조성할 수 있도록 임원의 의사 결정력 교육을 시행하고, 도전을

장려하고 실패를 두려워하지 않는 내부 분위기 조성 등의 노력이 필요할 것으로 보인다.

<표 10> 전체 그룹 '단점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
			사람	출근	퇴근	회사	영업	사원	신입	직원	퇴사	생각
1	근무환경	13.6%	11.9%	3.4%	3.4%	3.3%	2.7%	2.0%	1.9%	1.8%	1.4%	1.3%
2	보상체계	13.5%	10.2%	6.3%	3.7%	3.1%	3.1%	2.7%	2.2%	2.1%	2.1%	1.8%
3	성장 전략	12.6%	5.0%	2.6%	2.5%	2.4%	2.3%	1.9%	1.9%	1.7%	1.6%	1.4%
4	업무 강도	12.1%	15.7%	7.3%	5.9%	2.8%	1.8%	1.6%	1.5%	1.5%	1.3%	1.3%
5	조직문화	11.9%	13.6%	6.6%	6.2%	4.0%	3.5%	3.5%	3.2%	3.2%	3.0%	2.2%
6	평가체계	11.7%	4.2%	3.2%	1.9%	1.9%	1.9%	1.5%	1.3%	1.3%	1.3%	1.2%
7	휴가	9.4%	4.5%	4.3%	3.1%	2.6%	2.5%	2.3%	2.2%	2.1%	1.5%	1.5%
8	직급체계	7.8%	5.5%	5.5%	3.9%	3.5%	2.2%	2.2%	2.1%	2.0%	1.6%	1.6%
9	문서업무	7.4%	8.1%	6.3%	4.0%	3.9%	3.5%	3.1%	2.8%	2.8%	2.5%	2.1%

<표 11> 상 그룹 '단점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
1	업무강도	16.9%	업무 12.5%	부서 5.1%	근무 3.4%	강도 2.1%	경우 1.9%	체계 1.3%	환경 1.3%	현장 1.3%	차이 1.2%	야근 1.2%
2	보상체계	15.4%	연봉 8.8%	복지 5.0%	급여 2.9%	기업 2.8%	단점 2.6%	회사 2.3%	수준 2.0%	사원 1.8%	신입 1.7%	월급 1.7%
3	실적 압박	15%	회사 3.7%	영업 3.7%	직원 2.6%	임원 2.4%	실적 1.7%	압박 1.5%	매출 1.5%	구조 1.4%	경영진 1.4%	부족 1.3%
4	조직문화	14.4%	문화 11.8%	보수 5.2%	분위기 4.5%	기업 4.4%	조직 3.5%	수직 3.0%	회사 2.3%	군대 2.2%	꼰대 2.2%	군대식 1.4%
5	평가체계	11.3%	사람 16.4%	회사 3.6%	평가 2.0%	수당 1.9%	발전 1.8%	직원 1.8%	꼰대 1.7%	정치 1.6%	생각 1.6%	능력 1.6%
6	야근	10.4%	야근 4.0%	눈치 3.9%	퇴근 3.8%	출근 3.6%	연차 2.5%	상사 2.2%	분위기 1.8%	주말 1.7%	사용 1.6%	근무 1.6%
7	비정규직	9.6%	비정규 직 4.5%	정규직 3.4%	성장 2.8%	결정 2.6%	미래 2.5%	회사 2.5%	개발 2.4%	사업 2.3%	가능 2.0%	의사 1.9%
8	문서업무	7.2%	보고 10.4%	업무 5.0%	업계 4.9%	작업 4.2%	회의 3.7%	단순 3.5%	문서 3.4%	효율 2.8%	필요 2.4%	자료 2.3%

<표 12> 중 그룹 '단점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
			연봉	승진	인사	급여	복지	사람	신입	사원	진급	비정규 직
1	보상체계	13.4%	7.8%	3.4%	2.8%	2.5%	2.4%	2.3%	2.0%	2.0%	2.0%	1.9%
2	복리후생	12.6%	5.3%	5.1%	3.7%	1.9%	1.8%	1.7%	1.4%	1.4%	1.3%	1.3%
3	경영진	12.5%	5.7%	4.7%	3.7%	2.1%	2.1%	1.9%	1.8%	1.7%	1.7%	1.6%
4	실적 압박	12.3%	4.1%	2.4%	1.6%	1.6%	1.6%	1.5%	1.4%	1.4%	1.4%	1.3%
5	야근	11.8%	6.0%	4.5%	4.2%	4.1%	2.8%	2.7%	2.7%	2.7%	2.5%	2.2%
6	휴가 사용	11.2%	16.4%	6.8%	4.2%	2.8%	2.0%	1.9%	1.5%	1.4%	1.2%	1.2%
7	조직문화	11.2%	14.3%	5.5%	5.4%	5.4%	3.7%	3.3%	3.1%	2.3%	2.1%	2.0%
8	문서업무	7.9%	6.7%	5.0%	3.7%	3.0%	2.4%	2.4%	1.9%	1.6%	1.6%	1.2%
9	근무환경	7.1%	8.6%	7.2%	4.6%	3.2%	2.8%	2.5%	2.1%	1.6%	1.5%	1.3%

<표 13> 하 그룹 '단점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
1	업무체계	12.2%	업무 16.5%	부서 4.6%	보고 2.5%	체계 2.4%	효율 2.4%	강도 2.2%	작업 1.7%	시스템 1.4%	인원 1.4%	부족 1.3%
2	근무환경	12%	회사 3.5%	사람 3.4%	현장 1.7%	영업 1.4%	돈 1.2%	실적 1.2%	공장 1.1%	생각 1.0%	본사 0.9%	평가 0.8%
3	직급체계	11.7%	직원 4.3%	사원 4.1%	신입 3.8%	승진 3.0%	진급 2.7%	사람 2.3%	퇴사 2.2%	대리 1.8%	직급 1.8%	이상 1.7%
4	조직문화	11.6%	문화 11.7%	분위기 4.2%	보수 3.8%	폰대 3.3%	군대 3.3%	군대식 2.6%	기업 2.5%	수직 2.2%	회사 2.0%	사람 1.9%
5	성장전략	11.5%	회사 3.8%	사업 2.4%	영업 1.9%	성장 1.6%	제품 1.5%	투자 1.5%	발전 1.4%	가능 1.4%	단점 1.4%	미래 1.4%
6	임원의 의사결정	11.2%	회사 8.4%	직원 4.9%	구조 3.0%	사람 2.9%	임원 2.1%	결정 1.7%	분위기 1.7%	보수 1.7%	능력 1.7%	생각 1.6%
7	야근	11.1%	야근 5.5%	퇴근 5.1%	출근 4.4%	연차 3.9%	주말 2.9%	근무 2.8%	수당 2.7%	눈치 2.6%	사용 2.0%	강요 1.2%
8	보상체계	11.1%	연봉 10.9%	복지 4.9%	급여 4.6%	기업 3.1%	회사 2.3%	대비 2.2%	업계 2.2%	인상 2.2%	수준 1.9%	상승 1.7%
9	근무환경	7.5%	사람 7.9%	근무 6.0%	단점 2.2%	문제 1.8%	환경 1.7%	회사 1.5%	월급 1.4%	교대 1.4%	하루 1.2%	생산직 1.2%

(3) 기업 리뷰 '경영진에 바라는 점' 등급별 토픽 모델링 결과

'경영진에 바라는 점' 리뷰에 대해 전체 데이터 및 상, 중, 하 등급별로 구분하여 LDA 토픽모델링을 실시한 결과는 아래 <표 14> ~ <표 17>와 같다.

전체 리뷰에 대한 LDA 토픽 모델링 결과, '성장지원 체계', '합리적인 보상 체계', '인력 운영', '사업확대 및 투자', '조직문화', '효율 개선', '경영진 인식 개선', '인력 충원' 등이 주요 경영진에 바라는 점 토픽으로 도출되었음을 확인할 수 있다.

등급 별로 나타나는 특징을 보면, 가장 먼저 '상' 등급에서만 나타나는 특징으로 '비정규직 처우 개선'이 주요 토픽으로 언급되었음을 확인할 수 있다. '비정규직', '정규직', '대우', '전환' 등의 키워드가 함께 언급되며, 이는 '중', '하' 등급에서는 나타나지 않는 특징인 것을 확인할 수 있었다. 이러한 점에서 미루어 보아 '상' 등급에 해당하는 기업일 수록 비정규직 직원들이 정규직과의 대우 차이에 대해 불만을 느끼는 것으로 파악할 수 있으며, '정규직 전환 요구'로 리뷰 데이터에 나타나고 있음을 확인할 수 있다. 따라서, '상' 등급에 해당하는 기업은 비정규직 고용 이슈에 대해 제도적으로 문제는 없는지 다시 한번 점검하고, 비정규직 직원들과 정규직 직원들이 조화롭게 근무할 수 있는 노사관계 혁신 요소를 발굴해 나가야 할 것이다.

또한 '중' 그룹에서만 나타나는 특징으로는, '인사제도'가 주요 토픽으로 언급되었는데, '퇴사', '인사', '불만', '문제', '관리', '구조' 등의 키워드가 같은 토픽 내에서 함께 언급되는 점으로 보아 많은 직원들이 합리적인 성과 평가에 대해 문제의식을 가지고 있으며, 이는 이직 시도로까지 이어지고 있음을 시사한다. 따라서 '중' 그룹에 해당하는 기업들은 모든 구성원들에게 공정한 기회를 제공하는 인사제도 마련을 통해 긍정적인 직원 경험을 설계해 나가야 할 것이다(진주화, 2021). 이를 위해 기업에서는 구성원의 특징을 파악하기 위한

내부 임직원 설문조사 등을 통해 데이터를 수집하고, 이를 활용한 HR Analytics 를 통해 공정한 인사제도 설계가 이루어져야 할 것이다.

마지막으로, '하' 그룹에서만 나타나는 특징으로는 '경영진의 인식 개선' 토픽에서 '경영진', '퇴사', '마인드', '미래' 등의 키워드가 함께 언급되고 있는데, 이는 경영진이 임직원을 언제든지 교체할 수 있는 '소모품'으로 생각하며, 기존 임직원들이 퇴사하면 새로운 인원으로 충원하면 된다는 마음가짐을 지양하고, 직원을 회사의 '미래'로 인식하고 성장시키기 위한 경영진의 인식 개선을 요구하고 있음을 확인할 수 있다. 따라서, '하' 그룹의 기업에서는 임직원 교육 훈련 및 직무능력 개발을 위한 근로자 역량개발 프로그램 개발 및 지원이 이루어져야 할 것으로 보인다(한국 ESG 기준원, 2021).

<표 14> 전체 그룹 '경영진에 바라는 점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
			회사	직원	생각	발전	성장	퇴사	기업	경영진	인재	임원
1	성장지원 체계	17.9%	14.6%	8.6%	7.5%	3.5%	2.2%	2.2%	2.1%	1.8%	1.5%	1.2%
2	보상체계	15.4%	12.4%	10.6%	4.7%	4.1%	3.5%	2.9%	2.7%	2.6%	2.6%	2.5%
3	인력운영	14.2%	10.8%	8.9%	5.2%	2.8%	2.4%	2.2%	2.2%	2.1%	2.1%	2.0%
4	사업확대 및 투자	13.1%	8.6%	8.4%	4.6%	4.4%	4.1%	3.9%	2.9%	2.4%	2.0%	1.9%
5	조직문화	11.4%	10.3%	5.6%	4.8%	4.8%	3.2%	2.8%	2.3%	2.0%	2.0%	1.9%
6	합리적인 성과평가	10.7%	5.7%	3.7%	2.9%	2.5%	2.3%	1.8%	1.6%	1.6%	1.4%	1.2%
7	경영진 인식 개선	9.4%	5.2%	3.9%	3.2%	3.0%	3.0%	2.3%	1.7%	1.6%	1.5%	1.3%
8	인력 충원	7.9%	9.0%	6.1%	3.8%	비정규 직 3.5%	3.3%	3.1%	2.7%	1.8%	1.6%	1.4%

<표 15> 상 그룹 '경영진에 바라는 점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
			업무	필요	체계	개선	보상	복지	연봉	합리	시스템	인사
1	합리적인 성과평가	16.2%	6.6%	6.4%	4.5%	4.2%	2.9%	2.8%	2.6%	2.4%	2.4%	2.3%
2	조직문화	13.9%	8.6%	7.0%	5.6%	4.0%	2.8%	2.6%	2.6%	2.1%	1.9%	1.9%
3	복리후생	13.4%	27.8%	6.7%	4.5%	3.6%	2.4%	2.1%	1.6%	1.0%	1.0%	1.0%
4	경영진 소통	13%	7.0%	2.7%	2.4%	2.4%	2.4%	2.4%	2.2%	2.0%	1.8%	1.8%
5	사업확대 및 신규투자	13%	9.3%	8.4%	5.0%	4.6%	4.3%	3.2%	2.6%	2.5%	2.2%	1.6%
6	보상체계	10.8%	16.3%	4.9%	4.1%	3.2%	2.9%	2.2%	1.8%	1.7%	1.3%	1.2%
7	인력운영	9.9%	4.4%	3.8%	3.5%	3.5%	3.0%	2.8%	2.8%	2.7%	2.6%	2.1%
8	비정규직 처우개선	9.8%	5.9%	5.7%	4.1%	3.9%	3.4%	3.1%	2.1%	2.0%	1.8%	1.7%

<표 16> 중 그룹 '경영진에 바라는 점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
			회사	경영	성장	발전	필요	미래	직원	투자	기업	개발
1	성장지원	16.1%	11.5%	3.5%	3.3%	3.3%	2.5%	2.4%	2.3%	2.0%	2.0%	1.8%
2	복리후생	13.8%	10.5%	7.8%	7.7%	6.2%	5.0%	4.1%	3.4%	3.3%	3.0%	1.7%
3	업무체계	12%	5.9%	5.1%	4.2%	3.4%	3.2%	3.1%	2.9%	2.8%	2.8%	2.3%
4	인사제도	11.6%	7.7%	4.2%	3.9%	3.0%	2.7%	2.5%	2.4%	2.2%	2.1%	2.0%
5	경영진 인식개선	11.3%	11.4%	10.4%	5.9%	4.0%	2.3%	2.2%	1.6%	1.5%	1.5%	1.1%
6	인력운영	9.7%	9.4%	3.7%	3.6%	3.0%	비정규 직 3.0%	정규직 2.8%	채용 2.8%	인력 2.4%	효율 2.0%	기회 1.9%
7	보상체계	9%	21.3%	4.9%	4.5%	4.0%	3.7%	2.8%	2.7%	2.3%	1.4%	1.4%
8	사업확대 및 신규투자	8.9%	11.7%	8.5%	6.3%	5.5%	2.8%	2.5%	1.4%	1.2%	1.2%	1.1%
9	근무환경	7.6%	7.0%	6.2%	5.8%	2.6%	2.4%	2.3%	2.3%	1.6%	1.5%	1.4%

<표 17> 하 그룹 '경영진에 바라는 점' 리뷰 LDA 토픽모델링 결과

No	Topic	Prop	Keywords									
			직원	복지	회사	신경	급여	발전	기업	연봉	성장	개선
1	복리후생	15.9%	14.5%	11.4%	6.7%	5.4%	2.3%	1.7%	1.6%	1.6%	1.4%	1.3%
2	경영진 인식개선	15.2%	생각 11.4%	직원 9.5%	회사 8.3%	경영진 3.5%	경영 3.3%	퇴사 2.6%	마인드 1.8%	사람 1.7%	이유 1.3%	미래 1.3%
3	조직문화	12.5%	문화 8.2%	개선 5.8%	필요 4.2%	근무 3.9%	분위기 3.2%	환경 3.1%	현장 2.5%	조직 2.5%	소통 2.1%	기업 2.1%
4	사업확대 및 신규투자	11.6%	투자 8.6%	사업 7.1%	필요 5.1%	신규 4.0%	적극 3.7%	개발 3.1%	인원 2.3%	회사 1.8%	제품 1.5%	추진 1.5%
5	경영진 인식개선2	10.5%	사람 15.0%	돈 4.3%	회사 3.9%	문제 2.8%	생각 2.1%	경영진 1.9%	능력 1.7%	월급 1.6%	본인 0.9%	예전 0.9%
6	합리적인 성과평가	9.5%	체계 7.9%	필요 6.1%	보상 4.4%	시스템 4.3%	업무 3.9%	인사 3.7%	합리 3.5%	혁신 2.5%	평가 2.4%	전문 1.7%
7	인력운영	9.1%	사원 7.3%	신입 4.5%	영업 3.5%	회사 3.1%	교육 2.9%	채용 2.5%	필요 1.6%	업무 1.5%	결정 1.5%	임원 1.3%
8	보상체계	8.2%	연봉 5.8%	기업 3.2%	인재 2.0%	지금 1.7%	인상 1.7%	비정규 직 1.6%	연차 1.4%	상승 1.3%	업무 1.2%	방식 1.2%
9	효율 개선	7.5%	직원 8.3%	업무 5.6%	불만 4.5%	효율 3.8%	부서 3.3%	해결 2.8%	반영 1.8%	비전 1.7%	필요 1.4%	분배 1.4%

제 IV장 결론

제 1 절 연구 결과

본 연구에서는 ESG 등급에 따라 기업 리뷰 데이터에서 추출되는 토픽이 어떻게 다른지 파악하고, 기업의 사회 책임 경영을 통해 직무만족도를 제고하고 긍정적인 직원경험을 설계하기 위해서는 어떠한 노력들이 필요한지 연구하였다. 연구조사를 위하여 한국 ESG 기준원에서 평가한 2021 년 사회(S) 등급을 수집하여 등급별로 상, 중, 하 그룹을 나누었고, 잡플래닛에서 기업의 전,현직자가 작성한 기업 리뷰 데이터를 수집하여 그룹별로 LDA 토픽모델링을 적용하여 리뷰 데이터의 토픽을 각각 추출하였다. 주요 연구 결과는 다음과 같다.

첫째, '장점' 리뷰 분석 결과, 일과 삶의 균형을 뜻하는 '워라밸'이라는 토픽은 '상', '중' 그룹에서만 도출되었으며, '하'에서는 도출되지 않았다. 이는 '하' 그룹에 속하는 회사에서는 일과 삶의 균형을 유지할 수 있는 가족친화경영이 필요함을 시사한다. 또한, '상', '중' 그룹과는 달리 '하' 그룹에서는 '수평적인 조직문화'에 대한 언급이 없었음을 확인하였다. '하' 그룹에 속하는 회사에서는 조직문화 혁신 캠페인 등을 통해 수평적인 조직문화를 갖추기 위한 노력이 필요할 것으로 보인다. 마지막으로, '하' 그룹에서만 나타나는 특징으로 현장 근무에 대한 조건들이 자세히 나타나고 있음을 확인하였다. 이를 바탕으로 '하' 그룹에서는 안전보건 거버넌스 구축에 더욱 신경을 써야 할 것으로 보여진다.

둘째, '단점' 리뷰 분석 결과, '상', '중' 그룹에서는 '문서 업무(Paperwork)' 라는 토픽이 도출되었다. 이는 업무 효율성과 구성원의 근무 만족도를 저하시키는 요인으로 파악할 수 있는데, '상', '중' 그룹에 속하는 회사에서는 의사결정

단계를 간소화하고, 불필요한 문서 업무를 줄이는 Paperless 캠페인 시행 등이 필요할 것으로 보인다. 또한 '하' 그룹에서만 '임원의 의사결정' 토픽이 도출되었는데, 이는 변화를 두려워하고 책임을 회피하고자 하는 임원의 명확하지 못한 업무 지시와 의사결정이 단점으로 지적되고 있음을 의미하는 것을 확인하였다. 따라서 '하' 그룹에 속하는 회사에서는 임원의 의사결정력 교육 시행과 더불어 내부적으로 도전을 장려하는 분위기 조성이 필요할 것으로 보인다.

셋째, '경영진에 바라는 점' 분석 결과, '상' 그룹에서만 나타나는 특징으로 '비정규직' 토픽이 도출되었음을 확인하였다. '대우', '전환' 등의 키워드가 함께 언급된 것으로 보았을 때, 비정규직 고용에 근로기준법상 문제가 될 요소는 없는지 재점검이 필요하고, 비정규직이 정규직과 조화롭게 근무할 수 있는 노사관계 혁신이 필요할 것으로 보인다. 다음으로 '중' 그룹에서만 나타나는 특징으로, '인사제도' 토픽이 도출되었음을 확인하였다. 합리적인 성과 평가 시행에 대해 문제의식을 많이 가지고 있음을 확인하였는데, '중' 그룹에 속하는 회사에서는 모든 구성원들에게 공정한 기회를 부여하는 인사제도 마련을 통해 긍정적인 직원 경험을 설계해 나가는 것이 필요하며, 이를 위해서는 데이터를 기반으로 한 HR Analytics 가 수반되어야 할 것이다. 마지막으로, '하' 그룹에서는 '경영진의 인식개선'이라는 토픽이 도출되었다. 임직원을 언제든지 교체할 수 있는 '소모품'으로 인식하지 말고, 직원을 회사의 '미래'로 인식하고 교육 훈련 및 직무능력 개발을 위한 근로자 역량개발 프로그램 개발 및 지원이 필요할 것으로 보인다.

제 2 절 연구의 시사점

본 연구는 한국 ESG 기준원에서 평가한 2021 년 기업의 사회(S) 등급을 그룹별로 나누고, 그룹별로 잡플래닛 기업 리뷰(장점, 단점, 경영진에 바라는 점)를 LDA 토픽모델링 기법을 적용하여 분석하였다.

기존 연구에서는 잡플래닛 리뷰 데이터를 기업의 산업군을 기준으로 구분하여 분석하거나, 설문조사를 통해 임직원의 ESG 인식 수준을 조사하는 등 빅데이터를 활용한 ESG 경영 방안 제시에 대한 연구는 다소 미진하였다. 본 연구에서는 약 10 만 건에 달하는 잡플래닛 기업 리뷰 데이터를 분석하여 기업의 사회 책임 경영에 시사점을 제시했다는 점에서 기존 연구들과의 차별점을 가진다.

앞서 2 장의 연구문제에서 제시한 대로, S 등급별로 기업 리뷰(장점, 단점, 경영진에 바라는 점)를 분석하고 도출된 토픽은 등급별로 상이하게 나타나는 것을 확인할 수 있었다. 이는 기업의 사회 책임 경영 수준에 따라서 기업의 내부 관계자인 임직원의 이탈 방지와 긍정적인 직원경험 설계를 위해 추구해야 할 경영 가치가 다르다는 것을 의미한다.

본 연구는 S 등급을 상, 중, 하로 나누어 기업의 사회 책임 경영에 대해 각기 다른 시사점을 제시했다는 것과, 기업의 전/현직자가 남긴 잡플래닛 기업 리뷰를 활용해 기업 내부 이해관계자인 임직원 실제 의견을 반영하였다는 점에서 새로운 HR Analytics 의 연구방향을 제시하였다.

인적자원 운영 및 관리를 하는 기업 내 HR 부서와 ESG 경영 추진 전담 부서에서는 이러한 점에 착안하여, 소속된 회사의 사회 책임 경영 수준에 따라 인사제도 개선, 노사관계 혁신 및 긍정적인 조직문화 조성, 안전보건 거버넌스 구축 등 다양한 사회 책임 경영 계획을 보다 전략적이고 체계적으로 수립할 수 있다.

제 3 절 연구 한계점 및 제언

본 연구는 ESG 등급에 따른 기업 리뷰 데이터 토픽 모델링을 통해 기업의 사회 책임 경영에 대해 위와 같은 연구 결과를 도출하였지만, 다음과 같은 한계점이 존재하였으며 후속으로 이루어질 연구에 대한 방향을 다음과 같이 제시한다.

첫째, 한국 ESG 기준원에서 2021 년에 평가한 기업은 총 1,005 개였으나, 본 연구에서는 코스피 시가총액 기준 상위 404 개의 회사 리뷰 데이터만을 활용하였다. ESG 경영을 지속하는 기업은 계속해서 증가하는 추세이므로, 후속 연구에서는 더 많은 빅데이터를 활용해 볼 수 있다.

둘째, 본 연구에서는 2021 년의 ESG 평가등급을 활용하였으나, 기업 리뷰 데이터는 작성된 시기의 제한을 두지 않았다. 따라서 향후 연구에서는 데이터의 시계열 특성을 활용하여 연도별 ESG 등급과 그에 따른 리뷰 데이터에서 도출되는 토픽이 어떻게 변화하는지에 대해서도 확인할 수 있다.

셋째, 본 연구에서는 ESG 경영 중에서도 기업의 사회 책임 경영(S)에 대해서만 분석하였다. 경영진의 의사결정, 인식과 같이 지배구조(G)와 연계되는 분석 내용이 일부 있었으나, HR Analytics 라는 연구 특성상 인적자원 관리와 많은 연관이 있는 사회(S)의 비중이 대부분을 차지하게 되었다. 향후 연구에서는 환경(E)과 지배구조(G)와 기업 리뷰 데이터를 연계한 분석을 진행해볼 수 있다.

마지막으로, 본 연구에서는 LDA 토픽모델링 이후 토픽 간 연관성 분석 등 사후 분석이 이루어지지 않았다. 일관성(Coherence) 지표를 통해 최적의 토픽 수를 산출하였지만, LDA 토픽모델링 후 동시출현 단어분석과 같은 후속 검증이 이루어지지 못했다. 따라서 향후 연구에서는 LDA 토픽모델링 후 토픽 간 연관성 분석 등을 진행하는 것을 제언할 수 있다.

참고문헌

- 길호현(2018), “텍스트마이닝을 위한 한국어 불용어 목록 연구”, *우리말글*, 78, 1-25.
- 김동욱, 강주영, & 임재익(2016), “토픽모델링 기법을 활용한 산업별 직무만족요인비교 조사: 잡플래닛 리뷰를 중심으로”, *한국 IT 서비스학회지*, 15(3), 157-171.
- 김윤경(2020), “기업 비재무정보 (ESG) 공시가 재무성과와 기업가치에 미치는 영향”, *규제연구*, 29(1), 35-59.
- 김정수, 이석준(2016), “취업준비생 토픽 분석을 통한 취업난 원인의 재탐색”, *경영과 정보연구*, 35(1), 85-116.
- 김희은, 이한준(2021), “온라인 직원 리뷰와 기업가치”, *경영경제연구*, 43(2), 27-52.
- 박은정, 조성준(2014), “KoNLPy: 쉽고 간결한 한국어 정보처리 파이썬 패키지”, *제 26회 한글 및 한국어 정보처리 학술대회 논문집*, 133-136.
- 박재현, 한향원, 김나라(2022), “국내 ESG 연구동향 탐색: 2012~2021년 진행된 국내 학술연구 중심으로”, *벤처창업연구*, 17(1), 191-211.
- 백혜연, 박용석(2019), “기업 리뷰 웹 사이트 텍스트 분석을 통한 직원 불만 표현 추출과 불만 원인 도출 및 해소 방안”, *한국정보통신학회논문지*, 23(4), 357-364.ISO 690
- 석정미, 김민정, 이수열(2016), “택배 물류 서비스 종사자가 지각하는 기업의 사회적 책임활동이 조직몰입과 조직시민행동에 미치는 영향: 연령과 근속연수의 조절효과”, *서비스경영학회지*, 17(5), 151-173.
- 안병대, 신동원, 이한준(2022), “동기·위생이론에 근거한 직무만족 및 불만족요인에 관한 연구-토픽모델링과 잡플래닛 기업리뷰를 활용하여”, *경영과 정보연구*, 41(1), 39-56.
- 오상희(2021), “신용등급 및 ESG 등급이 기업가치에 미치는 영향에 관한 연구”, *세무회계연구*, 69, 125-144.
- 윤조현(2022), ESG 등급이 기업의 내부 이해관계자 관련 성과에 미치는 영향, 국내석사학위논문 이화여자대학교.
- 이재성, 김홍식(2010), “스마트워크 현황과 활성화 방안 연구”, *한국지역정보학회지*, 13(4), 75-96.

- 이재진. (2021. 04. 21). HR 리더의 전략적 선택, ESG. 월간인재경영, 2021년 4월호, 제194호. <http://www.abouthr.co.kr/news/articleView.html?idxno=4781>
- 이재진. (2022. 03. 02). ESG – 직원경험 – HR 애널리틱스로 이어지는 가치사슬(Value Chain). 월간인재경영 2022년 03월호, 제 205호.
<https://www.abouthr.co.kr/news/articleView.html?idxno=5165>
- 이정기, 이재혁(2020), ““지속가능경영” 연구의 현황 및 발전방향: ESG 평가지표를 중심으로”, *전략경영연구*, 23(2), 65-92.
- 이종서, 김성근, 강주영(2017), “직무 리뷰 분석을 통한 산업군별 직무만족/준속 요인 및 직무불만족/이직 요인에 관한 연구”, *한국 IT 서비스학회지*, 16(1), 1-26.
- 임옥빈(2019), “비재무적 정보가 기업성과에 미치는 영향: ESG 점수를 중심으로”, *국제회계연구*, 86, 119-144.
- 조현우(2022), “LDA 를 활용한 Batch 공정 모니터링 연구의 토픽 모델링 분석”, *한국산학기술학회 논문지*, 23(4), 26-33.
- 진주화. (2021. 07. 28). ESG 시대, HR의 역할 - Social(사회) 편. 월간인재경영, 2021년 7월호, 제198호.
<http://www.abouthr.co.kr/news/articleView.html?idxno=4929>
- 천민호, 설경진(2022), “호텔 직원의 ESG 활동 인식이 조직자부심과 조직유효성에 미치는 영향”, *관광레저연구*, 34(9), 161-175.
- 최기철, 이상용(2018), “직원을 위한 내부마케팅이 기업의 시가 총액 변동률에 미치는 영향 분석: 잡플래닛 기업 리뷰를 중심으로”, *Information Systems Review*, 20(2).
- 최진욱, 신동원, 이한준(2021), “IT 기업 직원의 만족 및 불만족 요인에 따른 이직률 예측: 토픽모델링과 머신러닝을 활용하여”, *한국데이터정보과학회지*, 32(5), 1035-1047.
- 한국ESG기준원(2021), ESG 모범기준
- 한승호, 김문섭, 서철승(2022), “ESG 활동이 직무만족에 미치는 영향: 지각된 기업이미지의 매개효과를 중심으로”, *한국경영공학회지*, 27(3), 61-72.
- 황용석(2007), “인터넷게시판 실명제에 대한 비판적 연구: 익명적 표현의 권리를 중심으로”, *언론과 사회*, 15(2), 97-130.
- Bailey, R., & Yeo, J. Jiang Lingjun, & Ferguson, A.(2020), ESG as a Workforce Strategy, Marsh &

McLennan Companies.

- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003), "Latent dirichlet allocation", *Journal of machine Learning research*, 3(Jan), 993-1022.
- Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J., & Blei, D. (2009), "Reading tea leaves: How humans interpret topic models", *Advances in neural information processing systems*, 22.
- Coughlan, M., P. Cronin, and F. Ryan (2009), "Survey Research : Process and Limitations", *International Journal of Therapy and Rehabilitation*, 16(1), 9-15
- Guo, Y., Barnes, S. J., & Jia, Q. (2017), "Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation", *Tourism management*, 59, 467-483.
- Jin, M., & Kim, B. (2022), "Effects of ESG Activity Recognition Factors on Innovative Organization Culture, Job Crafting, and Job Performance", *Administrative Sciences*, 12(4), 127.ISO 690
- Jung, Y., & Suh, Y. (2019), "Mining the voice of employees: A text mining approach to identifying and analyzing job satisfaction factors from online employee reviews", *Decision Support Systems*, 123, 113074.ISO 690
- Kang, H., & Yang, J. (2018), "Selection of the Optimal Morphological Analyzer for a Korean Word2vec Model", *Proceedings of the Korea Information Processing Society Conference* (pp. 376-379), Korea Information Processing Society.
- Kim, Jaeyun(2021), Text mining the trend of job satisfaction from employee's review through DMR topic model before and after COVID-19, 국내석사학위논문, Sungkyunkwan University
- Klotze, A(2022), "The Great Resignation Is Still Here, but Whether It Stays Is Up to Leaders."
- Kwon, H. J., Ban, H. J., Jun, J. K., & Kim, H. S. (2021), "Topic modeling and sentiment analysis of online review for airlines". *Information*, 12(2), 78.ISO 690
- Larkin, T. (2015), "Why you shouldn't ignore Glassdoor", *HR Magazine*, Sept 7 th, ISO 690
- Lee, J., & Kang, J. (2017), "A study on job satisfaction factors in retention and turnover groups using dominance analysis and LDA topic modeling with employee reviews on Glassdoor.com."
- Locke, E. A. (1976), The nature and causes of job satisfaction, *Handbook of industrial and organizational psychology*
- Morgan, J. (2017), "The employee experience advantage: How to win the war for talent by giving employees the workspaces they want, the tools they need, and a culture they can

- celebrate.”, John Wiley & Sons.
- Newman, D., Lau, J. H., Grieser, K., & Baldwin, T. (2010), “Automatic evaluation of topic coherence” *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics* (pp. 100-108).
- Plaskoff, J. (2017), “Employee experience: the new human resource management approach”, *Strategic HR Review*, 16(3), 136-141.ISO 690
- Seo, W., Oh, J., & Kim, Y. (2022), “The Effects of ESG Activities on Job Satisfaction, Organizational Trust, and Turnover Intention”, *인터넷전자상거래연구*, 22(1), 197-213.ISO 690
- Sievert, C., & Shirley, K. (2014), “LDAvis: A method for visualizing and interpreting topics”, *The workshop on interactive language learning, visualization, and interfaces* (pp. 63-70).ISO 690
- Sull, D., Sull, C., & Zweig, B.(2022), “Toxic culture is driving the great resignation”, *MIT Sloan Management Review*, 63(2), 1-9.
- TAKIMOTO, M. Y., & Srisuphaolarn, P. (2021), The influence of ESG on generation Z: a review of preference for workplace, Doctoral dissertation, Thammasat University
- Vijayarani, S., & Janani, R. (2016), “Text mining: open source tokenization tools-an analysis”, *Advanced Computational Intelligence: An International Journal (ACIJ)*, 3(1), 37-47.
- Won Joon Yoo, Introduction to Deep Learning for Natural Language Processing, Wikidocs

Abstract

ESG management value study using review data

- Focusing on corporate socially responsible management -

Lee, Seung Yup

Seoul School of Integrated Sciences and Technologies

Advisor: Oh, Tae Yeon

As ESG management has become a major strategy for corporate management as a non-financial value, interest in corporate social responsibility management has also increased. Accordingly, it has also become important to design positive employee experiences to executives and employees who are internal stakeholders of the company to provide job security and to manage human resources to prevent departure.

Domestic studies on the existing ESG grade mainly focused on estimating financial value through non-financial factors, and there was a limitation in that it was difficult to have representation due to the lack of a large number of samples mainly using surveys.

Therefore, the purpose of this study is to present practical implications for the company's social responsibility management strategy according to ESG management level by dividing the social (S) grade into groups and analyzing the review data of Jobplanet, which is increasing in value in the era of resignation.

The subjects of the study were the top 404 companies in the KOSPI market capitalization as of September 2022, and 104,452 social (S) grades and Jobplanet review

data of the company evaluated by the Korea Corporate Governance Service(KCGS) were text-minned and analyzed using tokenization and LDA topic modeling algorithms.

As a result of the study, first, in "The merit" review, the 'Work-Life Balance' topic was mentioned only in the upper and middle groups, and not in the lower groups. In addition, there was no mention of 'Horizontal Organizational Culture' in the lower group. Finally, in the lower group, it was confirmed that the conditions for field work appeared in detail. Through this, it suggests that companies belonging to the lower group need to improve family-friendly management and horizontal organizational culture that can ensure the balance of work and life of workers, and establish a safety and health governance system.

Second, in the "Disadvantage" review, the topic of 'Paper Work' was derived in the upper and middle groups. And, in the lower group, the topic of 'Executive Decision-Making' was derived as the main topic. Through this, the upper and middle groups need to improve work efficiency through the establishment of a smart work environment, and the lower groups need to enhance the decision-making power of executives and give appropriate authority.

Finally, in the review of "What they want from management", the topic of "Non-Regular Workers" was derived only from the upper group, and demands for the conversion of non-regular workers to regular workers and improvement in treatment were found. In addition, the topic of the 'Personnel system' was derived only from the middle group. Finally, only in the lower group, the topic of 'Improving Management's Awareness' was derived. Through this, the upper group needs to improve labor-management relations by improving the treatment of non-regular workers, and the middle group needs to prepare a personnel system that provides fair opportunities and evaluates all members. Finally, the lower group suggested that management's perception of employees as the future of the company needs to be improved to grow.

This study conducted HR Analytics, which designs positive employee experiences through online corporate review analysis, presents practical implications to prevent internal employees from leaving, and presents new research directions in that it differs by ESG management level of the company.

Key words: ESG, HR, HR Analytics, People Analytics, Job Review, Jobplanet, LDA,
Topic Modeling, Text Mining

Student Number: 2125418007