

※ 한국연구재단 평가에 참여하는 평가위원은 평가에서 알게 된 비밀을 다른 사람에게 누설하거나 직무상 목적 외 용도로 이용하여서는 안됩니다.(국가연구개발혁신법 제40조)

2022년 선정 신진연구자지원사업 결 과 보 고 서

사업명	2022년 선정 신진연구자지원사업					
과제번호	2022S1A5A8 057027	연구과제명	텍스트마이닝 방법론을 활용한 직무만족 결정요인 탐색과 실증 분석			
기본정보	연구책임자	주관기관	연구기간	연구비(간접비제외)		학생연구자
	이준혁	연세대학교 산학협력단	2022.05.01. ~ 2023.04.30. (1년)	1차년도	19,016천원	0명
				2차년도	천원	명
				3차년도	천원	명
				계	천원	명

※ 이를 위반할 경우 2년 이하의 징역 또는 2천만원 이하의 벌금에 처하게 됨을 안내드립니다.(국가연구개발혁신법 제41조)

목 차

I. 연구목적 및 방법	1
1. 당초 연구의 목적, 필요성 및 연구목표	1
2. 당초 연구 내용, 범위 및 방법	3
II. 연구수행 내용 및 결과	9
III. 연구결과 활용계획	17
1. 연구결과 활용계획	17
2. 연구성과	18
IV. 참고문헌	18

I. 연구 목적 및 방법

1. 당초 연구의 목적, 필요성 및 연구목표

오늘날 정보기술(ICT)의 비약적인 발전을 바탕으로 사회·경제 시스템 전반에서 디지털 전환(digital transformation)이 촉진되고 있다. 디지털 전환에 따라 생산·축적되는 데이터는 양적인 증대뿐만 아니라 사회·경제 전반에서 생산되어 다양성 또한 증가하고 있는 추세이다. 이러한 다양한 데이터는 조직 측면에서 새로운 비즈니스 기회를 창출하는데 활용될 수 있기 때문에 많은 조직들은 조직 내외에서 생산되는 양질의 데이터를 생산하거나 확보하기 위해 노력하고 있다(안창원·황승구, 2012; 이중학·Steven Kim·송지훈·채충일, 2020; Holsapple, Lee-Post & Pakath, 2014). 더욱이 데이터를 활용한 의사결정은 직감이나 경험에 입각한 의사결정의 모호성과 불확실성을 해소하고, 증거기반 경영(evidence-based management)을 실현하는데 기여할 수 있기 때문에 실무 분야에서 데이터 활용의 필요성이 지속적으로 강조되고 있다(양재완, 박우성, 2020; Robbins & Judge, 2011).

실무 분야뿐만 아니라 데이터 활용에 대한 중요성은 경영학 연구 분야에서도 강조되고 있으며, 이는 인적자원관리 및 조직행동 분야도 예외는 아니다(Luo, Zhou, & Shon, 2016). 그러나 아직까지 인적자원관리 및 조직행동 분야에서 데이터를 활용한 연구가 활발하게 진행되지 못하고 있다. 이러한 가장 큰 이유는 데이터 확보의 어려움 때문으로 볼 수 있다. 조직 및 기업에서 생산되는 HR 관련 데이터는 대부분 대외비로 취급되기 때문에 연구자들이 데이터에 접근하기 힘들고, 인터넷, 소셜미디어(예: 잡플래닛, 페이스북, 인스타그램)에 공개된 데이터를 수집하는 경우에도 대부분의 데이터가 비정형 데이터(unstructured-data)의 형태로 축적되어 있기 때문에 데이터를 수집하고, 가공하는데 여러 어려움이 존재한다.

이러한 어려움에 따라 아직까지 인적자원관리 및 조직행동 분야의 많은 연구들이 설문조사를 통해 수집된 연구 데이터를 기반으로 진행되어왔다. 그러나 설문조사를 통해 수집된 연구 데이터에는 부정적인 답변에 따라 발생할 수 있는 부정적인 편견이나 불이익을 방지하거나 조직과 상사에게 좋은 평판을 얻기 위해 응답자들이 거짓 또는 과장하여 바람직한 답변을 하게 되는 사회적 적절성 편견(social desirable bias)의 오류가 내포될 수 있다는 한계점이 있다(양재완, 박우성, 2020; Demaio, 1984; Fehr & Schmidt, 2006). 이는 보다 정확하고 현실적인 연구 결과를 도출하기 위해 설문조사 이외의 방법으로 수집된 데이터를 분석하여 기존의 연구결과와 비교해 볼 필요가 있다는 것을 의미한다.

따라서 본 연구에서는 기존의 연구들이 갖는 한계점과 데이터 활용의 어려움을 보완하기 위해 국내 기업 정보 플랫폼 중 하나인 잡플래닛(jobplenet.co.kr)을 통해 연구 데이터를 수집하여 연구에 활용하고자 한다. 잡플래닛은 사용자들(users)이 현

재 재직 또는 재직하였던 경험을 바탕으로 기업 및 직무환경의 장점과 단점을 자유롭게 기술한 자료와 직무에 대한 전반적인 만족도를 정량적으로 평가한 자료를 제공하고 있다. 특히, 기업 및 직무환경의 장·단점에 대한 질문은 개방형으로 구성되어 있기 때문에 연구자 개인의 입장에서 한정적으로 예측하고 측정하는 기존의 설문조사 방법 보다 구성원들이 가지고 있는 직무환경에 대한 다양한 만족(장점)과 불만족(단점) 요인을 심층적으로 살펴 볼 수 있는 장점이 있다(Luo et al., 2016).

이러한 맥락에서 최근 국내외 소수의 연구자들에 의해 기업 정보 플랫폼(예: jobplanet.co.kr, grassdoor.com)에 축적된 리뷰 데이터를 텍스트마이닝(text mining) 방법론을 활용하여 분석하려는 시도가 이루어지고 있다(김동욱, 강주영, 임재익, 2016; Jung & Suh. 2019; Luo et. al., 2016). 관련 연구들은 기존의 연구 데이터 수집과 관련된 한계점을 보완하고 새로운 연구방법론을 제시하였다는 점에서 높은 의의를 함양하고 있다. 그러나 구성원들의 기업 리뷰를 바탕으로 직무만족(job satisfaction) 또는 종업원만족(employee satisfaction)의 결정요인을 도출하는데 집중하여 실제 직무만족과의 인과관계를 다루지 못한 한계점이 있다.

이에 본 연구는 잡플래닛 사용자(user)가 작성한 기업 장점 리뷰(review) 데이터를 텍스트마이닝으로 분석하여 구성원들의 직무만족 결정요인을 도출하고, 도출된 요인들이 정량적으로 평가한 직무만족도에 미치는 영향을 함께 살펴보고자 한다. 이처럼 직무만족 결정요인을 도출하고 실제 직무만족도와의 인과관계를 살펴보는 이유는 도출된 각각의 요인들이 직무만족도에 매우 작은 영향을 미치거나 혹은 유의한 영향을 미치지 않을 수 있기 때문에 이를 검증할 필요가 있기 때문이다. 또한, 다중회귀분석을 통해 도출된 각 요인들이 직무만족도에 미치는 영향의 상대적인 크기를 비교할 수 있어 보다 효과적인 직무만족 향상 방안을 수립하는데 기여할 수 있다.

한편 본 연구는 기업 장점 리뷰 데이터와 함께 기업 단점 리뷰 데이터를 함께 분석에 활용하고자 한다. 이는 직무만족 결정요인의 수준이 낮은 경우 실제 직무만족이 감소 또는 유발되지 않거나, 혹은 직무불만족을 초래하는지 살펴보기 위함이다. Herzberg, Mausner & Snyderman(1959)은 2요인 이론을 통해 만족의 반대는 불만족이 아닌 만족하지 않는 상태라고 주장하면서, 직무만족을 유발하는 요인을 동기요인(예: 성취, 인정)과 직무불만족을 유발하는 요인을 위생요인(예: 정책, 관리)으로 구분하였다. 이는 직무만족 결정요인의 특성에 따라 해당 요인이 충족되었을 때 직무만족을 증가시키나, 결핍되어도 불만족을 유발하지 않는 요인이 있는 반면, 결핍되었을 때 불만족을 증가시키나 충족되어도 직무만족을 증가시키지 않는 요인이 존재할 수 있음을 의미한다. 따라서 이를 나누어 살펴보는 것은 상황에 따라 전략적으로 구성원들의 직무만족을 증가시키거나 직무불만족을 감소시키는 방안을 수립하는데 필요한 중요한 시사점을 제공할 수 있을 것이라고 판단된다.

본 연구는 이상의 목적을 바탕으로 기존의 설문조사 방법에 의존해오던 인적자원 관리 및 조직행동 연구 분야에 새로운 방법론을 제시하고, 구성원들의 보다 솔직한 응답을 바탕으로 실제 직무만족에 영향을 미치는 결정요인을 탐색함으로써 인적자

원관리 및 조직행동 분야에 다양한 이론적·실무적 시사점을 제시하고자 한다.

2. 당초 연구 내용, 범위 및 방법

2.1. 이론적 배경

2.1.1. 직무만족의 결정요인

직무만족은 구성원들이 자신의 직무에 대한 심리적 평가나 경험을 통해 느끼는 긍정적 감정 또는 이를 바탕으로 형성되는 정서적 태도를 의미한다(Locke, 1976; Hulin & Judge, 2003). 직무수행에서 발생하는 이러한 긍정적 감정과 태도는 직무를 수행하는 구성원 자신뿐만 아니라 조직에도 긍정적인 영향을 미치는 요인으로 인식되어, 조직행동 및 인사관리 연구 분야에서 조직유효성을 대표하는 주요 변인으로 활용되고 있다.

직무만족의 중요성 인식이 증가함에 따라 여러 연구자들에 의해 직무만족의 결정요인을 밝히기 위한 노력들이 이어지고 있으며, 대다수의 연구들에서 직무만족이 유발되는 원인을 구조적 또는 상황적 관점에서 접근하고 있다(방하남, 김상욱, 2009). 이러한 접근법은 직무에 대한 긍정적인 감정과 태도가 직무특성, 직무환경 등 직무의 구조나 상황에 의해 결정된다고 보는 관점이다. 예를 들어, Herzberg et al.(1959)은 2요인 이론(two factor theory)을 통해 직무환경 요인을 구성원들의 직무만족을 유발하는 동기요인(예: 성취, 인정)과 직무불만족을 유발하는 위생요인(예: 정책, 관리)으로 구분하여 제시하였으며, Hackman & Oldham(1975)은 직무특성이론(job characteristic theory)을 통해 구성원들의 직무만족을 증가시키는 5가지 직무특성(기술다양성, 과업정체성, 과업중요성, 자율성, 피드백)을 제시하였다.

그러나 구조 및 상황적 접근으로는 직무의 구조와 상황이 동일하더라도 왜 구성원들이 느끼는 직무만족 수준이 모두 동일하지 않은가를 설명하는데 한계가 있다. 이러한 차이는 구성원들이 자신의 가치관, 목적, 욕구 등을 토대로 객관적인 직무특성이나 직무환경을 주관적으로 평가하기 때문이다. 이는 직무만족 인식이나 수준이 개인특성에 따라 결정될 수 있음을 의미한다. 따라서 직무만족은 직무특성과 직무환경 뿐만 아니라 개인특성의 영향을 받으며, 각각의 특성들이 상호작용한 결과물로 볼 수 있다(Vroom, 1964; Locke, 1976; 방하남·김상욱, 2009).

이상의 논의를 바탕으로 많은 국내/외 선행연구들에서 직무특성, 직무환경 및 개인특성을 중심으로 다양한 직무만족의 결정요인이 제시되었다. 선행연구에서 제시된 대표적인 특성을 살펴보면, 먼저 직무특성으로 기술다양성, 과업정체성, 과업 중요성, 자율성, 피드백, 권한부여, 역할(모호성, 갈등, 명확성) 등이 제시되었다. 다음으로 대표적인 환경특성으로 임금, 복리후생, 경력개발, 고용형태, 승진기회, 고용안정

성, 노조가입 여부, 업무부하, 초과근무 여부, 교육수준 및 기술수준 적합도 등이 제시되었다. 마지막으로 대표적인 개인특성으로 인구통계학적 특성(나이, 성별, 학력, 혼인여부 등), 성격특성(개방성, 성실성, 외향성, 친화성, 신경증), 직무동기, 자존감 등이 제시되었다(Ali, Anis, & Yadav, 2015; Baker, 2004; Balouch & Hassan, 2014; Brown & Peterson, 1993; 방하남, 김상욱, 2009; 이홍직, 윤수인, 최순례, 윤승태, 2019).

이상의 연구들은 직무만족의 결정요인을 밝히는데 초점을 맞추고, 종합적인 관점에서 다양한 특성과 환경변수를 선행요인으로 고려함으로써 직무만족 결정요인을 다각적으로 규명하고 있다는 시사점에서 제공하고 있다. 그러나 연구설계에서 연구자가 직관에 의지하거나 선행연구를 바탕으로 자의적으로 직무만족 결정요인을 선정하고 데이터를 수집·분석한다는 점에서 한계가 있다. 이는 실제 현장에서 구성원들이 느끼고 인식하는 직무만족의 결정요인을 제대로 포함하지 못할 수 있음을 의미한다. 따라서 본 연구에서는 기존 연구의 한계점을 보완하기 위해 온라인 사이트(잡코리아)의 기업 리뷰 데이터(개방형 질문)를 활용하여 특정 변수에 한정하지 않고, 실제 현장에서 구성원들이 인식하는 직무만족의 결정요인을 탐색하고자 한다.

2.1.2. 텍스트마이닝

정보기술의 비약적인 발전과 소셜미디어의 확산을 바탕으로 텍스트, 이미지, 음성, 동영상 등 다양한 유형의 비정형 데이터가 폭발적으로 증가하고 있다. 이러한 비정형 데이터에는 의미 있는 정보들이 내포되어 있기 때문에 이를 분석함으로써 다양한 가치를 창출하고자 하는 노력이 증가하고 있다. 비정형 데이터 유형 중 가장 기본적이고 광범위한 비중을 차지하는 텍스트 데이터는 접근성과 활용도가 높아 많은 연구자들과 전문가들의 주목을 받고 있다.

텍스트 데이터는 자연어 처리(natural language processing) 기술에 기반을 둔 텍스트마이닝 기법을 통해 분석할 수 있다. 텍스트마이닝은 의미 있는 패턴이나 지식을 탐색하여 의미 있는 정보를 추출하거나 새로운 정보를 찾아내는 기법으로, 전처리 과정을 통해 문서 안의 비정형의 데이터를 정형화된 형태로 바꾸는 과정을 의미한다(서대호, 2019; Feldman & Sanger 2007; Hearst 2003).

텍스트마이닝의 절차는 크게 수집된 텍스트 데이터에 대한 자료처리와 자료분석 과정으로 구분된다(Hotho, Nürnberger, & Paaß, 2005). 먼저 자료처리 과정은 비정형화된 데이터를 자연어 처리를 통해 분석에 용이하도록 분류(classification) 및 전처리(pre-processing)하는 단계이다. 자료처리 과정은 일반적으로 (1) 문서의 정규화: 비정형 데이터를 정형 또는 반정형으로 변환, (2) 문서 분류: 정해진 기준에 따라 문서를 분류, (3) 문장 분리: 텍스트를 문장 단위로 분리, (4) 형태소 분석: 문장에서 의미를 가진 최소단위인 형태소로 분리, (5) 개체명 인식: 회사명, 인물명 등 하나의 개체로 인식 할 단어를 추출, (6) 불용어 제거: 분석결과에 의미 없는 단어를 제거 등의

절차가 포함된다(서대호, 2019; 전병진, 2019).

다음으로 자료분석 과정에서는 전처리된 자료를 텍스트마이닝 기법을 통해 분석하는 단계이다. 대표적인 분석 방법은 (1) 단어빈도분석: 문서에 출현한 단어의 빈도를 분석, (2) 동시출현분석: 문서에 두 개의 단어가 동시에 출현한 빈도를 분석, (3) 감성분석: 문장이나 단어가 가진 감성을 판별하고, 감성수준을 수치화, (4) 키워드 추출: 문서의 내용을 대표하는 키워드를 추출, (5) 군집 분석: 대량의 문서들을 서로 비슷한 성격 또는 주제로 묶어주는 분석, (6) LDA 토픽모델링: 군집 분석 기법 중 하나로 대량의 문서들에서 임의의 주제를 찾아내기 위해 유사한 의미를 가진 단어들을 여러 개의 그룹으로 나누어 분류 등이 있다(서대호, 2019; 전병진, 2019).

2.2. 연구의 추진전략 및 절차

앞서 ‘1. 연구의 필요성 및 목적’에서 제시한 본 연구의 목적을 달성하기 위해 다음 <그림 1>에 제시한 연구 추진도의 절차에 맞춰 체계적으로 연구를 진행하고자 한다.

1단계	2단계	3단계	4단계	5단계
1차 문헌조사 및 전문가 인터뷰	연구 데이터 수집	연구 데이터 분석	2차 문헌조사 및 전문가 인터뷰	연구 결과 논의 및 후속 연구 제언

<그림 1> 연구 추진도

2.2.1. 1차 문헌조사 및 전문가 인터뷰

1단계에서는 1차 문헌조사 및 전문가 인터뷰로 이를 통해 이론적 배경을 정리하고, 연구 데이터 수집 및 분석 방법을 구체화한다. 세부 절차는 다음과 같다. 첫째, 본 연구에서 살펴보고자 하는 직무만족 결정요인, 텍스트마이닝 방법론과 관련된 선행연구, 저서, 리뷰 등의 문헌자료를 토대로 이론적 배경을 정리한다. 이를 통해 직무만족과 결정요인, 텍스트마이닝의 정의, 연구동향 등을 정리하고, 본 연구의 목적을 바탕으로 조작적 정의 및 기존 연구와의 차별점을 구체화한다.

둘째, 교수, 연구자, 실무 전문가로 구성된 전문가와의 인터뷰를 통해 본 연구에서 수립한 연구변수의 조작적 정의와 본 연구의 차별성 및 독창성 등을 점검한다. 또한, 연구설계의 정교함을 높이기 위해 본 연구자가 수립한 연구 데이터 수집 및 분석 방법을 점검한다. 구체적으로 연구 데이터 수집을 위한 웹크롤링(web crawling)

방법, 수집할 데이터의 종류와 항목, 데이터 분석 절차 및 방법 등에 대해 논의하고 연구 진행에 발생 가능한 오류를 사전에 제거한다.

2.2.2. 연구 데이터 수집

2단계에서는 본 연구에서 사용할 기업 리뷰 데이터는 국내 구인, 구직, 기업 정보를 공유하는 소셜 미디어 사이트인 잡플래닛(www.jobplanet.co.kr)에서 수집하고자 한다. 잡플래닛에서는 12개의 산업군에서 유저들이 작성한 기업 장점, 단점 리뷰와 유저가 직접 정량적으로 평가한 직무만족도(5점 만점) 정보를 제공하고 있다. 연구자는 프로그래밍 언어인 파이썬(pyton)을 사용하여 해당 데이터를 웹크롤링으로 수집할 예정이다.

2.2.3. 연구 데이터 분석

3단계에서는 수집한 연구 데이터에 대한 분석을 진행한다. 본 연구에서는 비정형 데이터인 기업 리뷰와 정형 데이터인 직무만족도의 두 종류의 데이터를 수집할 예정이다. 이에 따라 본 연구에서 연구 데이터 분석을 크게 2단계 절차에 따라 분석하고자 한다. 1단계에서는 수집된 기업 리뷰 데이터에 대한 전처리와 텍스트마이닝 군집분석 기법인 LDA토픽모델링을 활용하여 직무만족 결정요인을 도출하고, 2단계에서는 도출된 직무만족 결정요인들이 직무만족도에 미치는 영향의 강도 및 유의성을 살펴보기 위해 다중회귀분석을 진행하고자 한다. 분석 방법 및 절차에 대해서는 다음 절인 ‘2.3. 연구의 방법 및 분석 절차’에서 보다 자세히 기술하고자 한다.

2.2.4. 2차 문헌조사 및 전문가 인터뷰

4단계에서는 2차 문헌조사 및 교수, 연구자, 실무 전문가로 구성된 전문가 인터뷰를 통해 도출된 연구 결과에 대한 논의를 진행한다. 구체적으로 본 연구는 기존의 설문조사 방법에 의한 연구와 다른 방법론을 선택하고 있기 때문에 문헌조사를 통해 선행연구의 결과들과 본 연구에서 도출된 결과의 유사점과 차이점을 비교하고 다양한 시사점을 도출한다. 또한, 전문가 인터뷰를 통해 분석결과에 대한 연구자의 해석에 오류가 없는지 검토하고, 본 연구자의 한정된 시각을 넘어 학계와 실무계의 다양한 관점을 반영함으로써 보다 효과적으로 구성원들의 직무만족을 증대시킬 수 있는 방안을 논의하고자 한다.

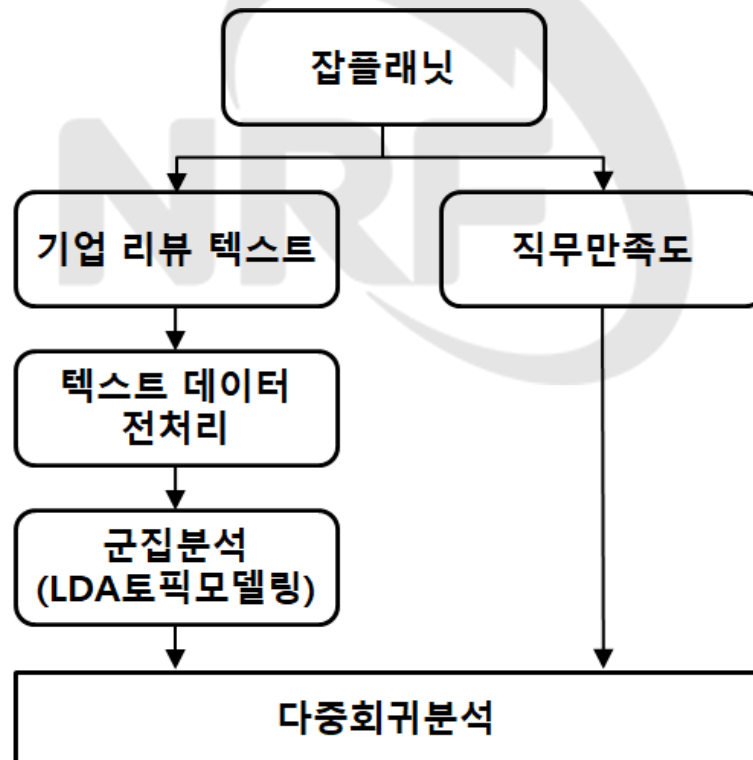
2.2.5. 연구 결과 논의 및 후속 연구 제언

마지막 5단계에서는 3단계에서 도출된 분석 결과를 정리하고, 4단계에서 조사·논

의된 연구 결과의 함의 및 이론적·실무적 시사점을 기술한다. 또한, 본 연구를 수행하면서 부족한 부분에 대한 한계점을 기술하고, 향후 연구에 대한 제언을 통해 후속 연구에 기여하고자 한다.

2.3. 연구의 방법 및 분석 절차

본 연구는 직무만족의 결정요인을 도출하고 실제 구성원들이 느끼는 직무만족에 어떠한 영향을 미치는지 밝히는 것을 목적으로 한다. 이러한 목적을 달성하기 위해 본 연구에서는 잡플래닛 온라인 사이트에서 파이썬 프로그래밍 언어를 이용하여 데이터를 수집하고 분석하고자 한다. 좀 더 구체적으로 본 연구는 1) 웹크롤링, 2) 텍스트 전처리, 3) 군집분석(LDA토픽모델링), 4) 다중회귀분석의 4가지 방법과 절차를 통해 수행될 예정이며, 이를 도식화하면 다음 <그림 2>와 같다.



<그림 2> 연구 방법 및 분석 절차

2.3.1. 데이터 수집: 웹크롤링

첫 번째 단계는 자료 수집을 위한 웹크롤링이다. 본 연구에 사용할 데이터는 온라인 웹페이지에 존재하는 대량의 텍스트 및 정량 데이터이기 때문에 이를 수집하기 위해서는 웹크롤링 기법을 사용하여야 한다. 웹크롤링은 웹페이지 또는 소셜 네트워크 서비스 내 게시판, 사용자 댓글, 문서 등에 존재하는 대량의 정보를 프로그래밍

언어를 활용해 반복적으로 수집할 수 있는 기법이다. 본 연구에서 자료를 수집하고자 하는 잡플래닛 사이트에는 12개의 산업군에 소속된 수천개 이상의 기업에 대한 수만개 이상의 방대한 리뷰와 평가가 존재한다. 따라서 본 연구에 활용할 데이터를 수집하기 위해서는 웹크롤링 기법을 사용할 필요가 있다.

2.3.2. 자료처리: 데이터 전처리

두 번째 단계는 텍스트 데이터 전처리이다. 본 연구에서 수집하는 데이터는 텍스트 데이터와 정량 데이터로 구분된다. 정량 데이터는 등간척도로 별도의 정제없이 바로 사용가능하지만 텍스트 데이터는 비정형 데이터이기 때문에 분석에 활용하기 위해서는 정형의 데이터로 변환하는 자료처리 과정이 필요하다. 따라서, 본 연구에서는 분석에 앞서 먼저 산업군, 기업 리뷰 장점과 단점으로 1) 문서를 분류하고, 각각의 기업 리뷰 장점과 단점 데이터를 정형화 시키기 위한 2) 문장분리, 3) 형태소 분석, 4) 개체명 인식, 5) 불용어 제거 등의 전처리 기법을 사용하고자 한다. 이에 따라 수행되는 각각의 전처리 기법 내용은 다음과 같다(서대호, 2019; 전병진, 2019; Feldman & Sanger 2007; Hearst 2003).

- 1) 문서 분류: 정해진 기준에 따라 문서를 분류하는 기법
- 2) 문장 분리: 텍스트를 문장 단위로 분리하는 기법
- 3) 형태소 분석: 문장에서 의미를 가진 최소단위인 형태소로 분리하는 기법
- 4) 개체명 인식: 회사명, 인물명 등 하나의 개체로 인식 할 단어를 추출하는 기법
- 5) 불용어 제거: 분석결과에 의미 없는 단어를 제거하는 기법

2.3.3. 자료분석: LDA토픽모델링

세 번째는 전처리된 데이터를 텍스트마이닝 기법을 통해 분석하고자 한다. 구체적으로 본 연구는 기업 리뷰 데이터를 통해 직무만족 결정요인을 도출하는 것을 목적으로 하기 때문에 텍스트마이닝 군집분석 기법 중 하나인 LDA토픽모델링을 사용하고자 한다. LDA토픽모델링은 대량의 텍스트를 요약하여 그 안에 내재된 의미를 찾아내는 기법으로 비구조화 되어 인간이 이해하기 어려운 대량의 텍스트 데이터를 인간이 이해하기 쉽게 구조화된 결과물로 도출할 수 있다는 장점이 있다. 또한, LDA토픽모델링은 확률적 분석을 바탕으로 대량의 텍스트 문서에서 유의미한 토픽과 키워드를 발견할 수 있으며, 토픽의 수와 토픽별 키워드의 수를 연구자가 직접 조정할 수 있어 토픽의 중복이나 의미 없는 토픽이 도출되는 것을 막을 수 있다(Blei 2012; 전병진 2019). 따라서, 본 연구에서는 LDA토픽모델링을 활용하여 앞서 전처리한 대량의 기업 리뷰 데이터 속에 내재되어 있는 유의미한 토픽과 키워드, 즉, 직무만족

의 결정요인을 도출하고자 한다. 또한, 본 연구는 기업 리뷰 장점과 단점에서 각각의 결정요인을 도출하여 비교하여 유사점과 차이점을 살펴보고자 하기 때문에 장점과 단점에서 직무만족 결정요인을 각각 도출하고자 한다.

2.3.4. 다중회귀분석

네 번째는 LDA토픽모델링을 통해 도출된 직무만족 결정요인이 실제 직무만족도에 미치는 영향을 분석하고자 한다. 구체적으로 본 연구는 LDA토픽모델링을 통해 기업 리뷰의 장점과 단점에서 각각 도출된 직무만족 결정요인들이 실제 직무만족도에 어떠한 영향(강도, 유의성)을 미치는지 검증하고자 한다. 따라서 회귀분석은 장점과 단점에서 도출된 결정요인으로 나눠서 2회 진행하고, 각 문서별로 도출된 전체 결정요인을 모두 투입하는 다중회귀분석으로 진행한다. 이처럼 다중회귀분석으로 2회 진행하는 이유는 다음과 같다.

먼저 장점과 단점에서 도출된 결정요인을 비교해 봄으로써 각 결정요인이 갖는 특성을 파악할 수 있다. 앞서 설명하였듯이 직무만족 결정요인 모두가 반드시 일차원적인 특성(직무만족을 증가시키는 요인이 부족한 경우 직무만족을 감소시킴)을 지닌다고 볼 수 없으며, Herzberg et al.(1959)이 2요인 이론을 통해 주장하였듯이 이차원적인 특성(직무만족을 유발하는 요인과 직무불만족을 유발하는 요인이 구분됨)을 갖는 요인도 존재할 수 있다. 결국, 이를 파악하기 위해서는 장점과 단점에서 각각 도출한 결정요인이 직무만족도에 미치는 영향을 비교해봄으로써 각 결정요인의 특성을 파악할 수 있을 것으로 판단된다.

다음으로, 다중회귀분석을 통해 직무만족 결정요인이 직무만족도에 미치는 영향을 검증하는 이유는 첫째, 다른 결정요인이 통제된 상태에서 특정 결정요인이 직무만족도에 미치는 순수한 영향을 살펴볼 수 있으며, 둘째, 도출된 각각의 직무만족 결정요인이 직무만족도에 미치는 영향의 상대적인 크기를 비교할 수 있기 때문이다. 따라서 다중회귀분석을 통해 직무만족 결정요인이 직무만족도에 미치는 영향을 살펴보는 것은 직무만족에 대한 각 결정요인의 중요도를 파악할 수 있고, 이를 바탕으로 직무만족을 효과적으로 향상시킬 수 있는 방안을 전략적으로 모색할 수 있다고 판단되기 때문이다.

II. 연구수행 내용 및 결과

1. 연구수행 내용

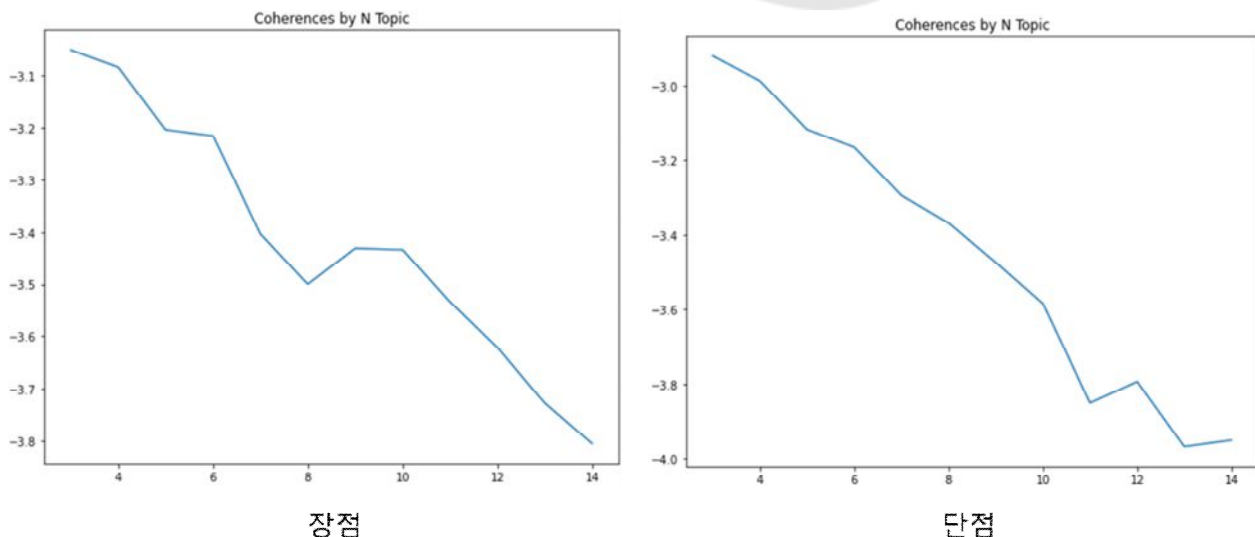
본 연구의 목적을 달성하기 위해 본 연구에서는 먼저 문헌조사와 전문가 인터뷰를 통해 연구 문제의 도출 및 분석방법에 대한 이론적·실무적 고도화 작업을 진행

하였다. 다음으로 도출된 연구 문제 검증을 위한 연구결과를 도출하기 위해 다음 4 단계의 체계적인 분석을 진행하였다. 첫째, 웹크롤링을 통해 총 271,170건(제조업 119,240건, 서비스업 68,176건, IT산업 83,754건)의 텍스트 데이터를 수집하고, 둘째, 수집된 텍스트 데이터 전처리를 통해 비정형 데이터를 정형화하였다. 셋째, LDA토픽 모델링을 통해 직무만족 및 불만족 결정요인 도출하였으며, 넷째, 다중회귀분석을 통해 결정요인과 직무만족도 간의 인과관계를 분석하였다. 마지막으로 도출된 연구 결과를 해석하고 심도 높은 시사점을 제시하기 위해, 추가적인 문헌조사 및 전문가 인터뷰를 통해 연구 결과의 해석과 시사점을 논의하였다.

2. 연구수행 결과

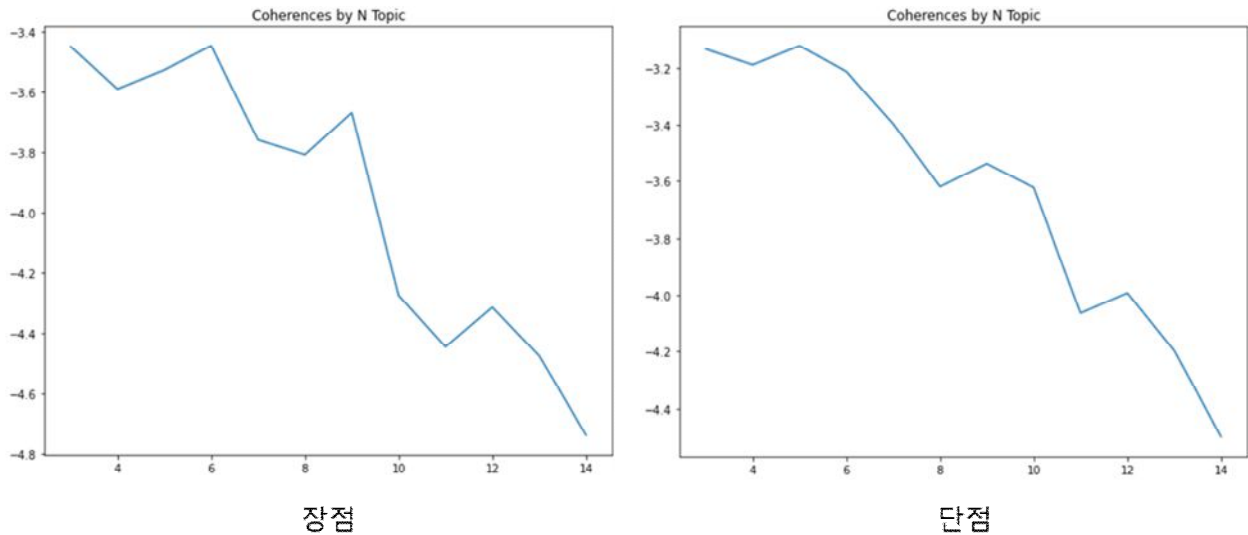
2.1. 토픽의 수 결정: 토픽 일관성 점수 분석 결과

토픽 모델링에 앞서 본 연구에서는 수집한 산업별 데이터 셋의 크기와 토픽 일관성 점수(coherence score)를 확인하여 적절한 토픽의 수를 결정하였다. 구체적으로 일관성 점수가 급격하게 감소하는 구간 이전의 토픽 수를 기준으로 하였으며, 이러한 기준을 바탕으로 제조업, 서비스업, IT산업의 장점과 단점 토픽 수 각각 6개를 도출하였다. 결정에 활용된 일관성 점수를 도식화 하면 다음 <그림 3>, <그림 4>, <그림 5>와 같다.

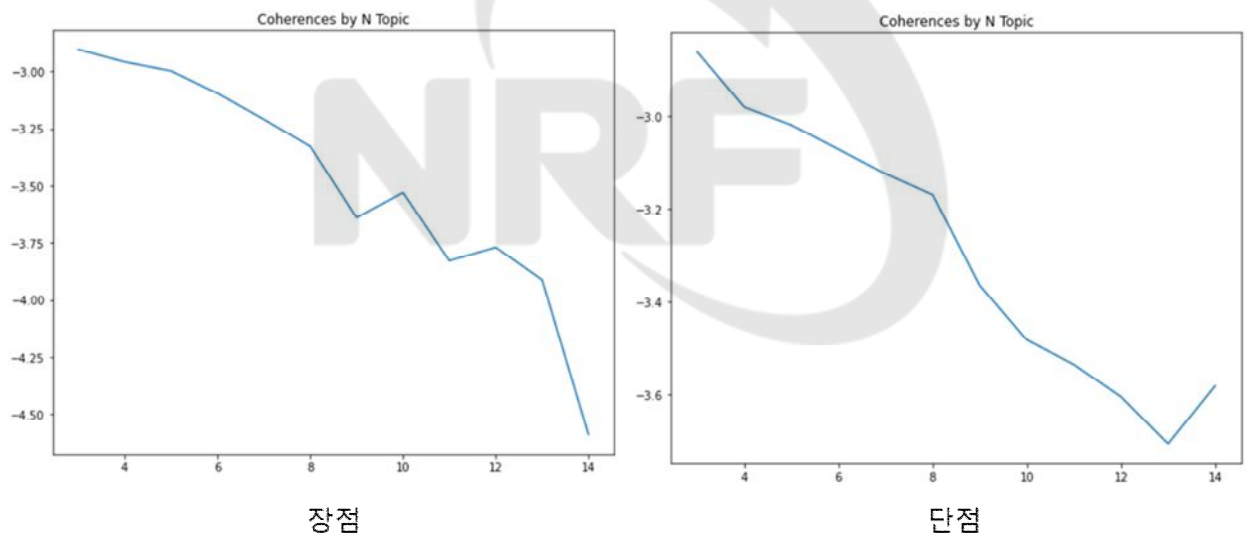


<그림 3> 제조업 토픽 일관성 점수

※ 한국연구재단 평가에 참여하는 평가위원은 평가에서 알게 된 비밀을 다른사람에게 누설하거나 직무상 목적 외 용도로 이용하여서는 안됩니다.(국가연구개발혁신법 제40조)



〈그림 4〉 서비스업 토픽 일관성 점수



〈그림 5〉 IT산업 토픽 일관성 점수

2.2. 산업별 장·단점 토픽모델링 분석 결과

앞서 결정된 산업별 장·단점 토픽 수에 따라 제조업, 서비스업, IT업을 구분하여 토픽모델링 분석을 실시하였으며, 구체적인 분석 결과는 다음과 같다.

2.2.1. 제조업

제조업 직무만족 및 불만족 결정요인을 도출하기 위해 제조업 근로자의 기업 장·단점 리뷰에 대한 LDA 토픽모델링을 수행하였다. 장점과 단점 리뷰를 각각 별도로 분석하였으며, 토픽 수는 일관성 점수 분석을 통해 결정된 6개를 적용하였다. 또한, 토픽별로 도출된 키워드 중 상위 또는 배타적으로 출현된 키워드를 중심으로

주요 키워드를 추출하고, 이를 바탕으로 각각의 토픽을 명명하였다. 구체적인 LDA 토픽모델링 분석결과는 다음 <표 1>과 같다.

먼저 장점 리뷰를 통해 도출된 직무만족 결정요인을 살펴보면, 토픽 1에는 문화, 분위기, 수평, 시스템, 사람, 가족 등 ‘수평적 조직문화와 구조’와 관련된 키워드가 포함되었다. 토픽 2에는 연봉, 대기업, 급여, 자부심, 혜택, 업계 등 ‘높은 급여수준’과 관련된 키워드가 포함되었다. 토픽 3에는 연차, 휴가, 눈치, 사육, 자율, 자유, 칼퇴 등 ‘일과 삶의 균형(work-life balance)’과 관련된 키워드가 포함되었다. 토픽 4에는 업무, 강도, 보장, 편하다, 쉬다 등 ‘낮은 업무강도’와 관련된 키워드가 포함되었다. 토픽 5에는 경험, 기회, 해외, 기술, 분야, 성장, 가능성 등 ‘높은 성장기회’와 관련된 키워드가 포함되었다. 토픽 6에는 복지, 기숙사, 시설, 버스, 할인, 식당, 식사 등 ‘복리후생’과 관련된 키워드가 포함되었다.

다음으로 단점 리뷰를 통해 도출된 직무불만족 결정요인을 살펴보면, 토픽 1에는 문화, 군대, 분위기, 수직, 나이, 경직, 꼰대 등 ‘수직적 조직문화와 구조’와 관련된 키워드가 포함되었다. 토픽 2에는 부서, 업무, 강도, 인사, 차이, 평가, 제도 등 ‘높은 업무강도와 비합리적 인사관리 체계’와 관련된 키워드가 포함되었다. 토픽 3에는 업무, 보고, 시스템, 효율, 회의, 관리, 보고서 등 ‘비효율적인 업무 프로세스’와 관련된 키워드가 포함되었다. 토픽 4에는 야근, 퇴근, 근무, 눈치, 주말, 연차, 특근 등 ‘일과 삶의 불균형’과 관련된 키워드가 포함되었다. 토픽 5에는 연봉, 대기업, 급여, 타기업, 급여, 월급, 수준 등 ‘낮은 급여수준’과 관련된 키워드가 포함되었다. 토픽 6에는 근무, 지방, 지역, 공장, 현장, 순환, 위치 등 ‘근무여건 및 환경’과 관련된 키워드가 포함되었다.

2.2.2. 서비스업

서비스업의 직무만족 및 불만족 결정요인을 도출하기 위해 서비스업 근로자의 기업 장·단점 리뷰에 대한 LDA 토픽모델링을 수행하였다. 분석 방법과 절차는 앞서 실시한 제조업과 동일하게 진행하였다. 이를 통해 도출된 구체적인 LDA 토픽모델링 분석결과는 다음과 같다.

먼저 장점 리뷰를 통해 도출된 직무만족 결정요인을 살펴보면, 토픽 1에는 사람, 직원, 회식, 동료, 가족, 음료, 나이 등 ‘대인관계’와 관련된 키워드가 포함되었다. 토픽 2에는 시간, 월급, 수당, 야근, 급여, 지급 등 ‘높은 급여수준’과 관련된 키워드가 포함되었다. 토픽 3에는 교육, 기회, 경험, 기술, 능력, 승진, 성장 등 ‘높은 성장기회’와 관련된 키워드가 포함되었다. 토픽 4에는 연차, 눈치, 사육, 자율, 휴가, 칼퇴, 야근 등 ‘일과 삶의 균형’과 관련된 키워드가 포함되었다. 토픽 5에는 업무, 근무, 환경, 수평, 문화, 개인 등 ‘수평적 조직문화 및 업무환경’과 관련된 키워드가 포함되었다. 토픽 6에는 복지, 할인, 혜택, 복리, 후생, 호텔, 포인트 등 ‘복리후생’과 관련된 키워드가 포함되었다.

다음으로 단점 리뷰를 통해 도출된 직무불만족 결정요인을 살펴보면, 토픽 1에는 직원, 계약, 평가, 인사, 제도, 능력, 실적 등 ‘비합리적 인사관리 체계’와 관련된 키워드가 포함되었다. 토픽 2에는 근무, 시간, 업무, 강도, 주말, 스케줄 등 ‘높은 업무강도’와 관련된 키워드가 포함되었다. 토픽 3에는 야근, 업무, 눈치, 퇴근, 연차, 주말, 출근 등 ‘일과 삶의 불균형’과 관련된 키워드가 포함되었다. 토픽 4에는 연봉, 문화, 분위기, 급여, 군대, 수직, 효율 등 ‘낮은 급여 수준과 수직적 조직문화’와 관련된 키워드가 포함되었다. 토픽 5에는 성수기, 매장, 여름, 겨울, 행사, 환경, 에어컨 등 ‘근무여건 및 환경’과 관련된 키워드가 포함되었다. 토픽 6에는 사람, 직원, 퇴사, 신입, 여자, 상사, 파견 등 ‘대인관계’와 관련된 키워드가 포함되었다.

2.2.3. IT업

IT산업의 직무만족 및 불만족 결정요인을 도출하기 위해 IT산업 근로자의 기업장·단점 리뷰에 대한 LDA 토픽모델링을 수행하였다. 분석 방법과 절차는 앞서 실시한 제조, 서비스업과 동일하게 진행하였다. 이를 통해 도출된 구체적인 LDA 토픽 모델링 분석결과는 다음과 같다.

먼저 장점 리뷰를 통해 도출된 직무만족 결정요인을 살펴보면, 토픽 1에는 연차, 눈치, 사용, 퇴근, 휴가, 칼퇴 등 ‘일과 삶의 균형’과 관련된 키워드가 포함되었다. 토픽 2에는 업무, 근무, 환경, 강도, 편함, 야근 등 ‘낮은 업무강도’와 관련된 키워드가 포함되었다. 토픽 3에는 지원, 복지, 카페, 점심, 식당, 저녁 등 ‘복리후생’과 관련된 키워드가 포함되었다. 토픽 4에는 경험, 교육, 프로젝트, 기회, 성장, 개발, 경력 등 ‘높은 성장기회’와 관련된 키워드가 포함되었다. 토픽 5에는 분위기, 문화, 수평, 복장, 구조, 자율, 존중 등 ‘수평적 조직문화 및 구조’와 관련된 키워드가 포함되었다. 토픽 6에는 연봉, 안정, 업계, 수준, 급여, 최고, 월급 등 ‘높은 급여수준’과 관련된 키워드가 포함되었다.

다음으로 단점 리뷰를 통해 도출된 직무불만족 결정요인을 살펴보면, 토픽 1에는 정치, 그룹, 사내, 정치가, 내부, 상황 등 ‘조직정치’와 관련된 키워드가 포함되었다. 토픽 2에는 연봉, 급여, 수준, 차이, 인상, 업계 등 ‘낮은 급여수준’과 관련된 키워드가 포함되었다. 토픽 3에는 사업, 프로젝트, 개발, 발전, 성장, 가능성 등 ‘낮은 성장기회’와 관련된 키워드가 포함되었다. 토픽 4에는 야근, 눈치, 퇴근, 근무, 연차, 휴가, 주말 등 ‘일과 삶의 불균형’과 관련된 키워드가 포함되었다. 토픽 5에는 직원, 계약, 정규직, 인력, 파견, 인턴 등 ‘비합리적 인사관리 체계’와 관련된 키워드가 포함되었다. 토픽 6에는 문화, 분위기, 보고, 수직, 제도, 군대, 구조 등 ‘수직적 조직문화 및 구조’와 관련된 키워드가 포함되었다.

2.3. 산업별 직무만족 및 불만족 결정요인과 직무만족도 간의 인과관계 분석

앞서 토픽모델링을 통해 도출된 직무만족 및 불만족 결정요인을 바탕으로 직무만족도에 미치는 인과관계를 분석하였다. 분석에는 더미변수를 활용한 다중회귀분석을 실시하였다. 그러나 분석결과를 종합적으로 살펴본 결과, 대다수의 토픽에서 유의성이 확보되지 않거나, 방향성이 일관적이지 않은 것으로 나타났다. 구체적인 분석 결과는 다음과 같다.

2.3.1. 제조업

제조업 직무만족 결정요인이 직무만족도에 미치는 영향을 분석한 결과는 다음 <표 1>과 같이 도출되었다. 구체적으로 살펴보면, 구성원들의 직무만족도에 수평적 조직문화 및 구조, 높은 급여수준, 일과 삶의 균형, 낮은 업무강도는 직무만족도에 유의한 정(+)의 영향을 미치는 반면, 복리후생은 유의한 부(-)의 영향을 미치는 것으로 나타났다. 그러나 높은 성장기회가 직무만족도에 미치는 영향은 유의하지 않았다.

<표 1> 제조업 직무만족 결정요인이 직무만족도에 미치는 영향

variable	coef	std err	t	P
const	2.9227	0.023	124.388	0
수평적 조직문화 및 구조	0.1061	0.039	2.719	0.007
높은 급여수준	0.1847	0.031	6.043	0
일과 삶의 균형	0.1126	0.034	3.359	0.001
낮은 업무강도	0.0733	0.033	2.214	0.027
높은 성장기회	0.0608	0.032	1.927	0.054
복리후생	-0.1085	0.042	-2.591	0.01
R-squared = 0.005, Adj. R-squared = 0.005, F-statistic = 12.23				

다음으로 직무불만족 결정요인이 직무만족도에 미치는 영향을 분석한 결과는 다음 <표 2>와 같이 도출되었다. 구체적으로 살펴보면, 구성원들의 직무만족도에 비효율적인 업무 프로세스와 낮은 급여수준이 유의한 부(-)의 영향을 미치는 것으로 나타났다. 그러나 수직적 조직문화 및 구조, 높은 업무강도 및 비합리적 인사관리 체계 일과 삶의 불균형, 근무여건 및 환경이 직무만족도에 미치는 영향은 유의하지 않았다.

〈표 2〉 제조업 직무불만족 결정요인이 직무만족도에 미치는 영향

variable	coef	std err	t	P
const	3.1457	0.021	150.196	0
수직적 조직문화 및 구조	0.0197	0.036	0.544	0.586
높은 업무강도 및 비합리적 인사관리 체계	0.0106	0.029	0.362	0.718
비효율적인 업무 프로세스	-0.0731	0.031	-2.389	0.017
일과 삶의 불균형	-0.0027	0.026	-0.103	0.918
낮은 급여수준	-0.0971	0.032	-3.019	0.003
근무여건 및 환경	-0.055	0.043	-1.266	0.205
R-squared = 0.001, Adj. R-squared = 0.001, F-statistic = 3.759				

2.3.2. 서비스업

서비스업 직무만족 결정요인이 직무만족도에 미치는 영향을 분석한 결과는 다음 〈표 3〉과 같이 도출되었다. 구체적으로 살펴보면, 구성원들의 직무만족도에 수평적 조직문화 및 구조, 높은 급여수준, 일과 삶의 균형, 낮은 업무강도는 직무만족도에 유의한 정(+)의 영향을 미치는 반면, 복리후생은 유의한 부(-)의 영향을 미치는 것으로 나타났다. 그러나 대인관계, 높은 급여수준, 높은 성장기회, 복리후생이 직무만족도에 미치는 영향은 유의하지 않았다.

〈표 3〉 서비스업 직무만족 결정요인이 직무만족도에 미치는 영향

variable	coef	std err	t	P
const	2.8132	0.028	100.26	0
대인관계	0.0203	0.04	0.511	0.609
높은 급여수준	0.0005	0.047	0.011	0.991
높은 성장기회	-0.001	0.039	-0.026	0.98
일과 삶의 균형	-0.0994	0.041	-2.443	0.015
수평적 조직문화와 낮은 업무강도	0.0879	0.039	2.248	0.025
복리후생	0.0807	0.042	1.941	0.052
R-squared = 0.003, Adj. R-squared = 0.002, F-statistic = 4.604				

다음으로 직무불만족 결정요인이 직무만족도에 미치는 영향을 분석한 결과는 다음

〈표 4〉와 같이 도출되었다. 구체적으로 살펴보면, 구성원들의 직무만족도에 비합리적 인사관리 체계는 유의한 정(+)의 영향을 미치는 것으로 나타난 반면, 낮은 급여수준과 수직적 조직문화는 유의한 부(-)의 영향을 미치는 것으로 나타났다. 그러나 높은 업무강도, 일과 삶의 불균형, 근무여건 및 환경, 대인관계가 직무만족도에 미치는 영향은 유의하지 않은 것으로 나타났다.

〈표 4〉 서비스업 직무불만족 결정요인이 직무만족도에 미치는 영향

variable	coef	std err	t	P
const	2.8395	0.028	100.352	0
비합리적 인사관리 체계	0.1637	0.036	4.505	0
높은 업무강도	-0.0415	0.026	-1.609	0.108
일과 삶의 불균형	-0.0485	0.027	-1.795	0.073
낮은 급여수준과 수직적 조직문화	-0.0748	0.027	-2.732	0.006
근무여건 및 환경	-0.0041	0.029	-0.141	0.888
대인관계	0.0275	0.029	0.937	0.349
R-squared = 0.003, Adj. R-squared = 0.003, F-statistic = 4.633				

2.3.2. IT업

IT업 직무만족 결정요인이 직무만족도에 미치는 영향을 분석한 결과는 다음 〈표 5〉와 같이 도출되었다. 구체적으로 살펴보면, 구성원들의 직무만족도에 높은 급여수준이 유의한 부(-)의 효과를 미치는 것으로 나타났다. 그러나 나머지 일과 삶의 균형, 낮은 업무강도, 복리후생, 성장기회, 수평적 조직문화 및 구조가 직무만족도에 미치는 영향은 유의하지 않았다.

〈표 5〉 IT업 직무만족 결정요인이 직무만족도에 미치는 영향

variable	coef	std err	t	P
const	2.976	0.035	84.385	0
일과 삶의 균형	0.0677	0.046	1.477	0.14
낮은 업무강도	-0.0502	0.062	-0.806	0.421
복리후생	-0.0004	0.053	-0.008	0.994
성장기회	-0.0791	0.069	-1.141	0.254
수평적 조직문화 및 구조	0.0455	0.061	0.749	0.454
높은 급여수준	-0.1286	0.062	-2.076	0.038
R-squared = 0.002, Adj. R-squared = 0.001, F-statistic = 2.158				

다음으로 직무불만족 결정요인이 직무만족도에 미치는 영향을 분석한 결과는 다음 <표 6>과 같이 도출되었다. 구체적으로 살펴보면, 구성원들의 직무만족도에 일과 삶의 불균형, 수직적 조직문화 및 구조는 유의한 부(-)의 영향을 미치는 것으로 나타났다. 그러나 조직정치, 낮은 급여수준, 낮은 성장기회, 비합리적 인사관리 체계가 직무만족도에 미치는 영향은 유의하지 않은 것으로 나타났다.

<표 6> IT업 직무불만족 결정요인이 직무만족도에 미치는 영향

variable	coef	std err	t	P
const	3.1148	0.031	100.804	0
조직정치	-0.101	0.057	-1.784	0.074
낮은 급여수준	-0.0393	0.053	-0.744	0.457
낮은 성장기회	-0.0422	0.055	-0.768	0.442
일과 삶의 불균형	-0.0922	0.046	-2.001	0.045
비합리적 인사관리 체계	0.0591	0.048	1.243	0.214
수직적 조직문화 및 구조	-0.0973	0.044	-2.229	0.026
R-squared = 0.003, Adj. R-squared = 0.001, F-statistic = 2.365				

Ⅲ. 연구결과 활용계획

1. 연구결과 활용계획

본 연구에서는 토픽모델링을 통해 제조업, 서비스업, IT업 종사자들의 직무만족 및 불만족 결정요인을 탐색하였으며, 산업별 각 6개의 직무만족 결정요인과 각 6개의 직무불만족 결정요인을 도출하였다. 나아가 도출된 직무만족 및 불만족 결정요인이 실제 구성원들의 직무만족도에 미치는 영향에 대한 더미변수를 활용한 다중회귀 분석을 시도하였다. 본 연구의 결과에서 다수의 직무만족 및 불만족 결정요인과 직무만족도에 미치는 영향이 유의하지 않고, 방향성이 일관적이지 않아 이를 활용하기에는 다소 어려움이 존재한다.

비록 본 연구에서 시도한 회귀분석의 결과를 활용하기 어려움이 있지만, 토픽모델링을 통해 산업별 직무만족 결정요인과 직무불만족 결정요인을 도출하였다는 점에서 본 연구의 결과는 다양한 의의를 갖는다. 특히 도출된 결과를 활용하여 직무만족 결정요인과 불만족 결정요인 간 비교, 산업별 직무만족 결정요인과 불만족 결정요인의 비교 분석이 가능하며, 이를 통해 조직이론 및 조직행동 분야에 유용한 이론적·실무적 시사점을 제공할 수 있을 것이다.

2. 연구성과

본 연구자는 이상에서 도출된 연구 결과에 대한 심도 높은 논의와 이론적·실무적 시사점을 포함하여, (사)대한경영학회에서 발간하는 연구재단 등재지인 대한경영학회지 36권 제 8호 (2023년 8월)에 게재하였다.

IV. 참고문헌

- 김동욱, 강주영, & 임재익. (2016). 토픽모델링 기법을 활용한 산업별 직무만족 요인비교 조사: 잡플래닛 리뷰를 중심으로. 한국 IT 서비스학회지, 15(3), 157-171.
- 방하남, & 김상욱. (2009). 직무만족도와 조직몰입도의 결정요인과 구조분석. 한국사회학, 43(1), 56-88.
- 서대호. (2019). 잡아라! 텍스트마이닝 with 파이썬. 서울: 비제이퍼블릭. 텍스트 마이닝을 활용한 ‘청소년상담, 1388, 2011-2018.
- 안창원, & 황승구. (2012). 빅 데이터 기술과 주요 이슈. 정보과학회지, 30(6), 10-17.
- 양재완, & 박우성. (2020). Workforce 애널리틱스의 이해와 전략적 활용. 조직과 인사관리연구, 44(2), 77-101.
- 이홍직, 윤수인, 최순례, & 윤승태. (2019). 임금노동자의 직무만족도에 영향을 미치는 요인에 관한 연구: 인구사회학적요인, 성격특성요인, 조직요인을 중심으로. 생명연구, 52, 85-107.
- 전병진. (2019). 소셜미디어 텍스트 마이닝을 위한 통합 애플리케이션 개발 : KoALA. 연세대학교 정보대학원. 석사학위논문.
- Ali, M., Anis, M., & Yadav, A. (2015). Antecedents of Job Satisfaction. International Journal of Economics, Commerce, and Management, 3(12), 350-360.
- Balouch, R., & Hassan, F. (2014). Determinants of job satisfaction and its impact on employee performance and turnover intentions. International journal of learning and development, 4(2), 120-140.
- Blei, D. M. (2012). Probabilistic topic models. Communications of the ACM, 55(4), 77-84.
- Brown, S. P., & Peterson, R. A. (1993). Antecedents and consequences of

- salesperson job satisfaction: Meta-analysis and assessment of causal effects. Journal of marketing research, 30(1), 63-77.
- DeMaio, T. J. (1984). Social desirability and survey. Surveying subjective phenomena, 2, 257.
- Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism-experimental evidence and new theories. Handbook of the economics of giving, altruism and reciprocity, 1, 615-691.
- Feldman, R., & Sanger, J. (2007). The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge university press.
- Hackman, J. R., & Oldham, G. R. (1975). Development of the job diagnostic survey. Journal of Applied psychology, 60(2), 159.
- Hearst, M. (2003). What is text mining. SIMS, UC Berkeley, 5.
- Herzberg, F. (1959). Mausner. SNYDERMAN., The motivation to work, New York.
- Holsapple, C., Lee-Post, A., & Pakath, R. (2014). A unified foundation for business analytics. Decision support systems, 64, 130-141.
- Hulin, C. L., & Judge, T. A. (2003). Job attitudes.
- 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.
- Luo, N., Zhou, Y., & Shon, J. (2016). Employee satisfaction and corporate performance: Mining employee reviews on glassdoor. com.
- Vroom, V. H. (1964). Work and motivation.