



2023년 제25회 융합학술대회



한국경영정보학회 & 한국생산관리학회
융합세션 중

한국경영정보학회 발표 논문집

- 일시 : 2023.8.16(수) 13:00 ~ 16:55
- 장소 : 부산 BEXCO
제1전시장 2층 214호 - Session B1
제2전시장 1층 122호·123호 - Session A1/A2/B2



2023년 융합학회 한국경영정보학회 세부 프로그램

Session A

BEXCO 제2전시장 1층 122호

A1 지속가능한 경영과 ESG 좌장 : 정의범 (한신대)

1	산업별 지속가능경영 전략 분석 : ESG 보고서와 뉴스 기사 분석	김원희, 권영옥(숙명여대)	2
2	사이버 위협 대응 모의훈련 통합형 평가체계 개발	박진용, 이용수, 김태성(충북대)	12
3	EU 공급망 실사법 대응을 위한 중소기업 ESG 평가지표 개선방안 연구	김수지, 기준학(숙명여대)	13
4	텍스트마이닝을 활용한 기업의 ESG 중대성(Materiality) 평가 정합성 분석 연구	김해리, 기준학(숙명여대)	23

A2 서비스와 플랫폼 좌장 : 이지은 (한양사이버대)

5	라이브 커머스 쇼핑이 소비자 구매의도에 미치는 영향	리우주, 차오신, 오세환(경북대)	36
6	The Price Formation of Emerging Digital Art Market : Timing and Market Experience of NFT users	김우경, 정동일(숙명여대)	45
7	건강 기능성 게임 수용에 영향을 주는 요인에 관한 연구	김용영(건국대)	53
8	네트워크 분석을 활용한 대학 전공 분석 : 12대 국가전략기술을 중심으로	소정은, 박병조, 김태성(충북대)	62

Session B

BEXCO 제1전시장 2층 214호

B1 빅데이터와 경영인사이트 좌장 : 강주영 (아주대)

9	개인간 연합학습 방법을 적용한 혈압 관리 시스템 : Digital Me 알고리즘을 기반으로	김수현, 정백, 이경전(경희대)	64
10	AUC 최적화 학습 알고리즘의 기업 부실 예측 연구	안재현, 김명중(부산대)	65
11	조건부가치측정법을 이용한 비식별화된 이미지 가치 측정	최유정, 김태성(충북대)	70
12	금융 범주 불균형 최적화를 위한 AUC 부스팅 학습	권성훈, 김명중(부산대)	71
13	MSA 정보화 사업의 정보시스템 감리 방안에 대한 제언	김종원, 박진형(한국과학기술정보연구원)	76
14	마진 분포 최적화를 통한 부실예측 모형의 예측 성과 개선	조성임, 김명중(부산대)	88

2023년 융합학회 한국경영정보학회 세부 프로그램

Session A

BEXCO 제2전시장 1층 122호

A1

지속가능한 경영과 ESG

좌장 : 정의범 (한신대)

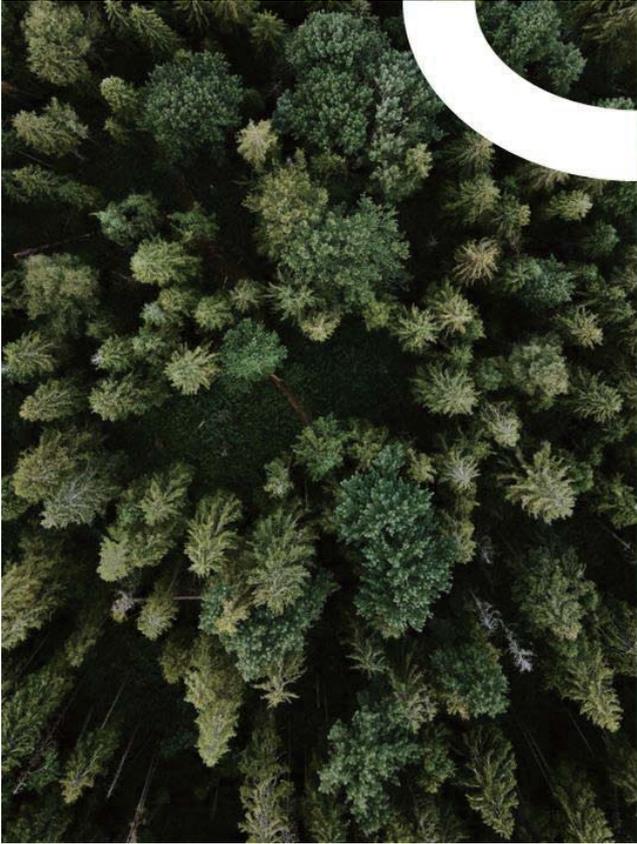
- | | | |
|---|------------------------------------------------------|--------------------|
| 1 | 산업별 지속가능경영 전략 분석
: ESG 보고서와 뉴스 기사 분석 | 김원희, 권영옥(숙명여대) |
| 2 | 사이버 위협 대응 모의훈련 통합형 평가체계 개발 | 박진용, 이용수, 김태성(충북대) |
| 3 | EU 공급망 실사법 대응을 위한 중소기업 ESG 평가지표
개선방안 연구 | 김수지, 기준학(숙명여대) |
| 4 | 텍스트마이닝을 활용한 기업의 ESG 중대성(Materiality)
평가 정합성 분석 연구 | 김해리, 기준학(숙명여대) |

A2

서비스와 플랫폼

좌장 : 이지은 (한양사이버대)

- | | | |
|---|---------------------------------------------------------------------------------------------------|--------------------|
| 5 | 라이브 커머스 쇼핑이 소비자 구매의도에 미치는 영향 | 리우주, 차오신, 오세환(경북대) |
| 6 | The Price Formation of Emerging Digital Art Market
: Timing and Market Experience of NFT users | 김우경, 정동일(숙명여대) |
| 7 | 건강 기능성 게임 수용에 영향을 주는 요인에 관한 연구 | 김용영(건국대) |
| 8 | 네트워크 분석을 활용한 대학 전공 분석
: 12대 국가전략기술을 중심으로 | 소정은, 박병조, 김태성(충북대) |



산업별 지속가능경영 전략 분석 : ESG 보고서와 뉴스 기사 분석

A Study on Analysis of Industry-specific Sustainability Strategy
: ESG Reporting and News Article Analysis

김원희, 권영옥

숙명여자대학교 빅데이터분석융합학과, 경영학부

이메일: {won2.96, yokwon}@sookmyung.ac.kr

◆◆◆ 목차

1. 서론	4. 연구 결과
연구 배경 3	ESG 보고서 분석 결과 13-14
연구 목적 4	뉴스 기사 분석 결과 15-16
연구 문제 5	
2. 선행 연구	5. 결론
ESG 관련 연구 6	연구 결과 정리 17-19
텍스트 마이닝을 활용한 ESG 연구 7	
3. 연구 방법	
분석 대상 선정 8	
데이터 수집 9	
데이터 전처리 10	
분석 방법 11-12	

01 서론

◆ 연구 배경

• ESG

기업의 비재무적 성과 평가요소인 환경(environment), 사회(social), 지배구조(governance)를 총칭하는 용어로 지속가능경영과 사회적 책임(CSR)의 규범화된 개념



환경(Environment)



사회(Social)



지배구조(Governance)

[그림 1] ESG

- 2006년 ESG를 투자 결정 및 자산에 고려한다는 UNPRI(유엔책임투자원칙)이 발표되고, 2017년 TCFD의 기후변화 관련 재무 정보공시 가이드라인이 발표됨에 따라 지속가능성을 고려한 기업경영 강조
- 국내외 경우 한국기업지배구조원(KCGS)이 2011년부터 ESG항목을 통해 국내 상장기업의 지속가능경영 수준을 평가하고 있으며, 2021년 한국거래소의 기업공시제도 개선안에 따르면 자율 공시였던 기업의 지속가능경영 보고서 발간이 2025년부터 자산 2조원 이상의 기업에게 의무화하고 2030년부터는 모든 코스피 상장사로 확대됨에 따라 ESG 경영 전략 수립 필요
- ESG 요소를 고려한 지속가능경영의 중요성이 높아짐에 따라 기업의 ESG 경영 전략 수립을 위한 지속가능경영 트렌드 분석 필요

3

01 서론

◆ 연구 목적

• 산업별 ESG 경영 전략

- 비즈니스의 특성에 따라 환경, 사회, 거버넌스 이슈의 영향도와 관련성이 다르고 기업별 산업 특성 및 조직 상황에 맞는 전략 수립이 중요
- 글로벌 표준기구에서는 기업의 ESG 관련 이슈 *중대성 평가에도 산업별 다른 샘플을 제공, 국내의 ESG 평가기관에서는 업종의 특성에 따라 다른 가중치를 적용하여 ESG 점수를 부여 *중대성 평가: 기업 경영환경과 이해관계자들의 평가에 따라 기업의 주요 이슈를 평가하는 것
- 산업 특성에 따라 다른 ESG 관련 이슈의 중대성 평가가 수립되며 때문에 산업별 차별화된 ESG 경영 전략을 세우는 것이 중요

• ESG 세부항목 별 중요도에 따른 산업 분류

- 환경 요소 중요도가 높은 산업: 금융기업에 의해 '환경 위험 관리 대상'으로 정의된 산업
- 사회 요소 중요도가 높은 산업: KCGS 평가 기준을 통해 사회적 책임 경영 이슈가 드러난 산업
- 지배구조 요소 중요도가 높은 산업: 글로벌 ESG 정보공시 기준인 SASB 기준에서 강조되는 산업
- 각 세부항목 별 중요도에 따라 '제조분야', 'IT 분야', '금융 분야'를 분석 대상 산업 분야로 선정

환경(E)	사회(S)	지배구조(G)
농업, 임업, 어업, ...코크스·연탄 및 석유정제품 제조업, 금속 가공제품 제조업, ... 제조 분야	에너지, 소재, 자본재, ... 소프트웨어와 서비스, 기술 하드웨어와 장비, ... IT 분야	자산관리·보관업, 소비자금융, 보험업, 증권·상품 거래업... 금융 분야

<표 1> ESG 세부항목 별 중요도가 높은 산업 분류

4

01 서론

◆ 연구 문제

• 연구문제 1

- ESG 보고서와 뉴스 기사 텍스트 분석을 통해 도출된 산업별 ESG 경영 전략 및 ESG 요소별 핵심 이슈는 무엇인가?

• 연구문제 2

- ESG 보고서를 통해 공시하는 경영 전략과 뉴스 기사에 나타나는 ESG 이슈는 어떠한 관계성이 있는가?

• 연구문제 3

- 산업별 ESG 경영 전략과 트렌드의 종합적 탐구를 통해 얻을 수 있는 전략적 시사점은 무엇인가?

5

02 선행 연구

◆ ESG 관련 연구

1. ESG가 기업의 재무 성과에 미치는 영향을 분석하는 연구

Corporate Sustainability: First Evidence on Materiality (Mozaffar Khan, George Serafeim, and Aaron Yoon, 2014)

- 기업의 지속가능성 투자와 재무적 성과 간의 상관관계를 분석한 연구로 기업의 재무적·비재무적 지속가능 요인에 대한 투자 행위 모두 재무적 성과에 유의한 양(+)의 영향을 미치는 것을 확인

ESG 활동의 효과와 기업의 재무적 특성 (강원, 정무권, 2020)

- ESG 활동이 기업가치에 미치는 효과가 기업의 재무적 특성에 따라 어떤 차이를 보이는지 탐구하여 ESG 활동의 질적인 요소들이 기업가치에 미치는 영향에 대한 회귀분석 및 강건성 분석을 통해 상관관계 파악

2. 기업 가치와 ESG 사이의 관계를 분석하는 연구

Enhancing Market Valuation of ESG Performance: Is Integrated Reporting Keeping its Promise? (Mervelskemper and Streit, 2017)

- ESG성과 공시 여부와 ESG 보고서 공시 형태가 기업 가치에 어떠한 영향을 주는지 연구하며 회계 기반 가치평가 모형을 적용해 ESG 성과의 기업 가치 연관성 분석

비재무적 정보가 기업성과에 미치는 영향 : ESG 점수를 중심으로 (임옥빈, 2019)

- 비재무적 정보인 기업의 사회적 책임이 기업성과에 미치는 영향을 분석하며 한국기업지배구조원의 ESG등급 평가 지수 및 점수를 관심변수로 이용하고 기업가치 측정 변수로 Tobin's Q를 사용하여 기업의 세부적인 경영전략이 기업가치에 미치는 영향 확인

→ 기업 가치에 긍정적으로 작용하는 지속가능경영 전략 수립 방향성을 제시하고, 세부적으로 ESG 요소별 경영 전략을 탐구하는 연구 필요

6

02 선행 연구

◆ 텍스트 마이닝을 활용한 ESG 연구

1. 텍스트 마이닝을 활용해 ESG 연구동향을 탐색하는 연구

국내 ESG 연구동향 탐색: 2012~2021년 진행된 국내 학술연구 중심으로 (박재현, 한향원, 김나라, 2022)

- ESG 관련 국내 학술연구에 대한 텍스트 마이닝을 통해 국내 ESG 연구동향을 탐색하며 2012년부터 2021년까지의 국내 ESG 관련 학술 논문을 수집하고, 주제어와 논문 제목을 활용해 키워드 빈도분석 진행

텍스트 마이닝을 통한 ESG 연구동향 분석 (이가은, 최영준, 2023)

- 국제 학술지에 투고된 ESG 관련 논문을 분석해 ESG에 관한 국제 연구동향을 파악하며 2013년부터 2022년까지 투고된 논문 초록을 수집해 키워드 빈도분석, 토픽 모델링, 중심성 및 네트워크분석을 수행하고 기간별 분석 결과 해석

2. 특정 산업 분야에서의 ESG 연구

텍스트 마이닝을 활용한 금융업 세부 업종간 ESG 보고서 비교 분석 (박수빈, 2018)

- 금융업을 은행, 보험, 증권 3개 업종으로 나누어 ESG 활동 우수 기업들의 경영 전략 분석 및 정책 시사점 논의하며 ESG 활동 우수 업종의 ESG 보고서를 대상으로 키워드 빈도 분석과 의미연결망 분석 수행

텍스트마이닝 기법을 활용한 SNS 상에서 항공사 친환경 ESG경영인식이 기업명성, 태도 및 고객시민행동에 미치는 영향 (고민환, 2023)

- 항공사의 ESG 경영에 대한 대중의 인식이 기업명성, 태도, 고객시민행동에 미치는 영향 관계 탐색하며, '항공사 ESG 경영 환경'을 키워드로 포털에 등장한 기사를 수집하여 CONCOR 분석을 통해 대중 인식을 파악하고 실증연구를 통해 상관관계 탐구

→ 선두적으로 ESG를 실천하는 기업의 경영 전략과 트렌드를 파악할 필요가 있으며 다양한 산업 분야에서의 ESG 경영 활동 연구 필요

03 연구 방법

◆ 분석 대상 선정

1. ESG 평가등급 우수기업

분석 대상으로 국내 주요 ESG평가사인 한국ESG기준원(KCGS)과 서스틴베스트 평가등급을 종합하여 산업별 상위 2개 회사를 선정

기업명	KCGS				서스틴베스트		
	ESG통합	환경	사회	지배구조	전체등급	규모등급	
금융 분야	KB금융그룹	A+	A+	A+	A+	AA	AA
	신한금융그룹	A+	A+	A+	A+	AA	AA
제조 분야	SK	A+	A	A	A+	평가제외	
	포스코	A	A	A	A+	A	A
IT 분야	네이버	A	A	A	A	AA	A
	카카오	A	A	A	A	A	BB

<표 2> 기업별 ESG 평가등급

- 금융 분야, 제조 분야, IT 분야의 세가지 산업 중 국내 ESG 평가기관 산정 최고등급인 A+(KCGS), AA(서스틴베스트)를 획득한 기업을 우선적으로 선정
- SK주식회사는 자체 사업을 진행하지 않는 순수 지주회사기 때문에 서스틴베스트의 평가대상에서 제외되었으며 KCGS 등급을 반영

03 연구 방법

◆ 데이터 수집

1. ESG 보고서

분석 대상 기업 6개의 ESG 보고서를 각 사 홈페이지에서 수집, 보고서 내에서 불필요한 그림, 표, Appendix를 제외한 텍스트만 사용

수집 방법

ESG 보고서는 모두 국제 지속가능성 보고 기준인 GRI Standards의 핵심 요건(Core Option)에 부합하게 작성되었기 때문에 각 기업의 GRI Index를 활용해 전체 보고서를 4개의 파트*로 나누어 수집

*Overall Strategy, Environment, Social, Governance의 4개 파트로 구분

2. 뉴스 기사

뉴스 빅데이터 검색엔진인 빅카인즈에서 각 분석 대상 기업의 ESG에 관한 보도 기사를 수집 및 활용

수집 방법

- ① 전국일간지, 경제일간지 등 지역별 54개 언론사의 보도기사 내에서 수집 대상 '기업명'과 'ESG'를 모두 포함하는 기사를 검색하여 수집
- ② 수집된 기사에 대한 직접 검토를 통해 최종 분석 대상으로 수집된 기사 활용

	수집기간	
	ESG 보고서	뉴스 기사
금융 분야	2019 ~ 2021년 보고서	2019.01.01 ~ 2022.12.31
제조 분야	2019 ~ 2021년 보고서	2019.01.01 ~ 2022.12.31
IT 분야	2020 ~ 2021년 보고서	2019.01.01 ~ 2022.12.31

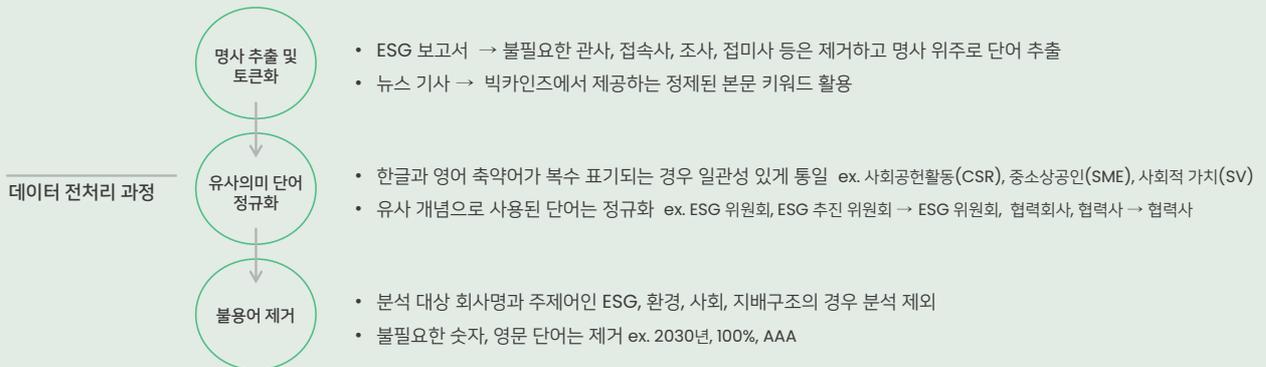
<표 3> 데이터 수집 기간

03 연구 방법

◆ 데이터 전처리

1. 분석 도구

수집한 데이터에 대한 전처리와 분석에는 Python 3.9를 활용하며 한글 텍스트 처리에는 KoNLP(Korean Natural Language Processing)패키지, LDA 토픽모델링 수행에는 Gensim 패키지 이용



03 연구 방법

◆ 분석 방법

1. 키워드 빈도 분석

분석 대상 텍스트 내에서 자주 등장하는 단어를 추출하여 문서 내 빈도를 분석하는 기법으로 데이터 내의 핵심 키워드를 빠르게 파악 가능

- ESG 보고서
 - 분석 대상 기업인 'KB금융그룹', '신한금융그룹', 'SK', '포스코', '네이버', '카카오'를 산업별로 통합하여 분석에 활용
 - 산업·파트별로 분류된 빈도 분석 결과를 종합하여 공통 어휘와 단독적으로 사용된 어휘 파악
- ESG 관련 기사
 - 분석 대상 기업인 'KB금융그룹', '신한금융그룹', 'SK', '포스코', '네이버', '카카오'를 산업별로 통합하여 분석에 활용
 - 산업별로 공통적으로 도출된 상위 빈도 키워드 비교·분석

산업명	기업명
금융 분야	KB금융그룹, 신한금융그룹
제조 분야	SK, 포스코
IT 분야	네이버, 카카오

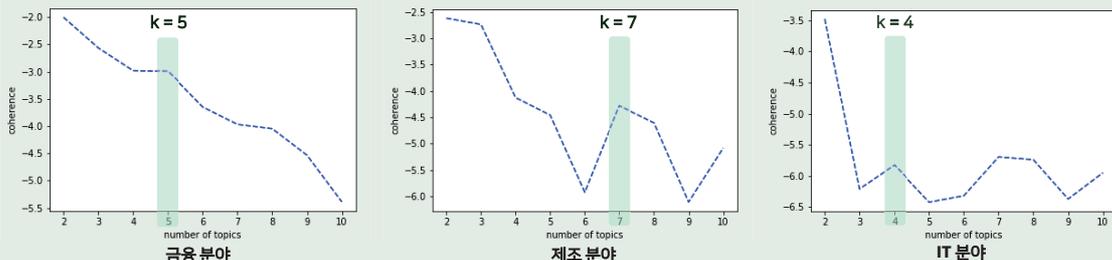
<표 4> 분석 대상 기업 산업 분류

03 연구 방법

◆ 분석 방법

2. LDA 기반 토픽 모델링

- 문서 집합에서 문서들이 어떤 주제로 이루어졌는지 파악하는 기법이며 도출된 토픽을 통해 비슷한 토픽을 가진 문서들을 묶어서 분석 가능
- 잠재 디리클레 할당 모델(Latent Dirichlet Allocation, LDA)은 대표적인 토픽 모델링 기법 중 하나로 단어들의 확률을 이용하여 문서집합 내의 잠재된 토픽들을 찾아내는 기법
- 토픽수 선정은 *응집도(Coherence)와 **혼잡도(perplexity)를 지표로 활용하여 통계적으로 적절한 개수를 결정하는 사후 지정 방식을 선택
 - *응집도: 문서 내에서 단어 간 연결의 정도를 평가하는데 사용되는 지표로 응집도가 커질수록 연결성이 높아졌다고 판단
 - **혼잡도: 엔트로피를 모델링한 것으로 혼잡도가 낮을수록 토픽이 안정적으로 구성되어 무질서도가 낮은 상태를 의미
- 본 연구에서는 뉴스 기사 분석 시 2~10개 토픽을 비교하여 토픽의 중복이 없이 유의미한 결과를 도출한다고 판단되는 토픽 수를 결정

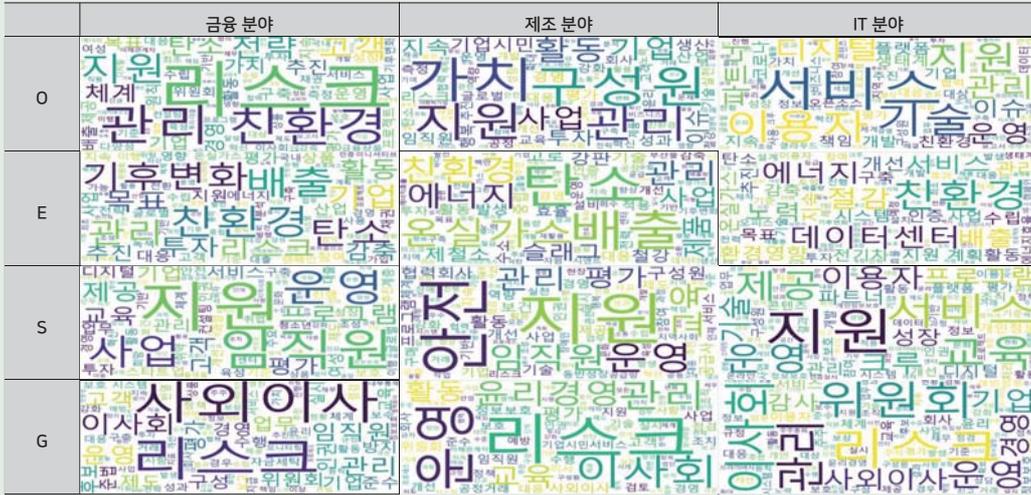


<표 5> 산업별 토픽 개수에 따른 응집도 그래프

04 연구 결과

◆ ESG 보고서 분석 결과

1. 키워드 빈도 분석 결과 - 워드 클라우드



<표 6> 산업별·파트 별 빈도 기반 워드 클라우드

04 연구 결과

◆ ESG 보고서 분석 결과

1. 키워드 빈도 분석 결과 - 산업별·ESG 요소별 주요 키워드 분류

	금융 분야	제조 분야	IT 분야
O	고객, 전략, 탄소, 체계, 목표, 구축, 배출, 제공, 원칙	구성원, 투자, 기업, 시민, 임직원, 생산, 성과, 행복, 산업, 평가, 글로벌	서비스, 기술, 디지털, 파트너, 플랫폼, 코로나, 생태계, 개발, 책임, 오픈소스
	가치, 기업, 지원		
E	기후변화, 기업, 리스크, 투자, 평가, 상품, 산업, 채권	제품, 슬래그, 제철소, 강판, 철강, 고로, 기술, 효율, 발생, 수소, 중립	데이터센터, 절감, 서비스, 노력, 환경영향, 개선, 지속, 수립, 전기차, 시스템
	관리, 배출, 에너지, 친환경, 활동		
S	고객, 여성, 업무, 투자, 스타트업, 인권, 컨설팅	활동, 협력회사, 보건, 공급사, 역량, 제도	이용자, 크루, 성장, 파트너, 개인정보, 플랫폼, 정보, 콘텐츠
	관리, 교육, 운영, 제공, 지원, 평가, 프로그램		
G	업무, 구성, 후보, 제도, 인권, 준수, 방지, 준법	윤리경영, 활동, 사업, 지원, 점검, 공정거래, 기술	감사, 서비스, 윤리, 체계, 주주, 개인정보, 규정
	관리, 교육, 리스크, 사외이사, 운영, 이사회, 위원회		

<표 7> 산업별·파트별 주요 키워드 분류

04 연구 결과

◆ 뉴스 기사 분석 결과

1. 키워드 빈도 분석 결과 - 금융 분야, 제조 분야, IT 분야

금융 분야 빈도수 순위					
키워드	n	키워드	n	키워드	n
지원	2292	에너지	524	추진	395
사업	1651	관리	522	인쇄	389
고객	1403	뉴딜	507	신설	389
전략	1281	평가	504	정보	387
글로벌	1244	기후변화	499	조직	385
서비스	1108	탄소중립	493	선언	383
플랫폼	1043	미래	492	개회	377
계획	979	협업	491	마사회	376
성장	827	스타트업	475	역량	376
지속가능	707	캠페인	471	금융사	371
상품	682	협약	464	전환	368
시장	680	디지털	462	기관	367
책임	674	관계자	460	정책	363
공영	671	육성	456	협력	356
규모	653	가입	452	가치	353
선정	645	대표	450	경제	353
혁신	642	개발사	441	상생	351
대출	614	성과	441	사회이사	344
노력	585	선도	431	변화	334
자산	575	교육	429	구성	331
프로그램	575	코로나19	428	연말	328
채권	561	성출	426	공동	328
프로젝트	550	탄소배출량	420	발전	328
업무	546	데이터	405	직원	326
발행	529	감독	397	운용	323

<표 8> 금융 분야 키워드 빈도 순위

제조 분야 빈도수 순위							
키워드	n	키워드	n	키워드	n	키워드	n
사업	5070	운영	855	기여	634		
경연	3498	노력	840	비즈니스	631		
글로벌	2626	안전	823	관계사	630		
탄소중립	1821	설명	817	RE100	624		
생산	1802	논의	812	발굴	623		
전략	1673	성장	811	설립	619		
기업시민	1656	개선	803	책임	618		
소재	1543	추진	798	상비	616		
협력	1406	공급	784	플랫폼	614		
평가	1399	전환	768	공장	606		
혁신	1338	투자	765	소통	606		
미국	1277	전문	747	행복	604		
세터	1188	지속가능	742	협력사	602		
철강	1085	관리	732	넷제로	599		
구성원	1059	신설	730	연말	597		
협력	1025	경쟁력	711	발전	589		
방안	1011	솔루션	696	장부	584		
생태계	994	전력	694	위원회	580		
사업	965	인사	693	조직	578		
경영	961	코로나19	671	인프라	572		
경제	920	프로그램	660	신뢰	569		
반도체	887	개회	656	공정	566		
재생에너지	885	대표	650	체제	565		
역량	876	구축	636	친환경	561		
발전	873	프로젝트	634	기갑까지	557		

<표 9> 제조 분야 키워드 빈도 순위

IT 분야 빈도수 순위					
키워드	n	키워드	n	키워드	n
서비스	1405	이용자	374	인사	250
디지털	864	윤리	369	지역	246
경영	790	글로벌	349	인터넷	242
플랫폼	748	스타트업	346	지속	239
인공지능	684	개선	337	생태계	239
지원	668	활동	328	방안	237
책임	587	상생	325	설명	236
계획	569	재생에너지	322	콘텐츠	236
공동	568	보고서	305	시장	231
데이터센터	555	초성	300	소상공인	230
운영	551	프로젝트	298	매출	229
성장	520	지속가능	296	코로나19	227
공동체	463	연말	296	인프라	222
협력	461	정보	289	관리	218
교육	451	선언	287	RE100	214
감화	444	정책	287	해결	211
사업	441	데이터	282	에너지	211
투자	437	프로그램	282	채권	211
참여	431	조직	278	발행	210
노력	420	전통시장	272	채널	203
전략	400	펀드	269	신인	203
혁신	393	신설	264	수입	202
위원회	389	업권	261	실업	198
미래	389	가치	260	대표	196
이사회	375	온라인	258	비즈니스	195

<표 10> IT 분야 키워드 빈도 순위

04 연구 결과

◆ 뉴스 기사 분석 결과

2. LDA 토픽 모델링 결과 - 금융 분야, 제조 분야, IT 분야

토픽 1	토픽 2	토픽 3	토픽 4	토픽 5
사업	고객	전략	탄소중립	서비스
경영	성장	글로벌	소비자	협약
육성	미래	자산	고객	에너지
글로벌	플랫폼	감축	기후변화	플랫폼
컨설팅	직원	지속가능	센터	고객
규모	소통	탄소배출량	협약	채권
소호사관학교	고객중심	선정	공동	시장
청년	문화	선언	생활	계획
지영업자	전략	가입	탄소배출량	전략
성장	조직	여성	공헌	공동
스퀘어브릿지	관리	책임	기부	건축
혁신	교육	지수	아동	인수
계획	디지털	노력	쓰레기	경영
채권	사업	평가	소비	협업
발행	내실	수입	업무	협력
뉴딜	경제	탄소중립	전달	계약
신청	서비스	대출	지원	상품
전략	책임	탄소	계획	거래

<표 11> 금융 분야 토픽 모델링 결과

토픽 1	토픽 2	토픽 3	토픽 4	토픽 5	토픽 6	토픽 7
글로벌	재생에너지	생산	기업시민	구성원	협력사	채권
배터리	RE100	철강	성장	방안	복구	투자
미국	글로벌	제품	가치	행복	포항철소	시장
협력	달성	공급	경쟁이념	소통	공급사	글로벌
전략	평가	탄소중립	전환	토론	폐쇄	에너지
혁신	주주	원료	추진	신뢰	개선	미국
생산	전략	발생	탄소중립	논의	프로그램	성공
산업	위원회	인증	전략	아이디어	동반성장	지속가능
전문	자회사	건설	성과	코로나19	지원	관리
솔루션	온실가스	설비	참여	공감	상생	투자자
역량	반도체	채널	이차전지	혁신	교육	본드
설립	탄소중립	추진	투자	역할	감소	자금
생태계	분할	협력	소재	개회	구매	관심
소재	유류제	구축	이사회	온라인	펀드	목적
베트남	노력	공정	포항	위기	공정	소재
바이오	주주총회	전기차	변화	협의회	평가	기준
조직	넷제로	저감	참석	개념	매출	철강회사
포트폴리오	안전	공정	체제	사람	운영	성장
비즈니스	선언	협약	안전	전략	안전	조달
아시아	설명	소재	발굴	기업가치	스타트업	망기

<표 12> 제조 분야 토픽 모델링 결과

토픽 1	토픽 2	토픽 3	토픽 4
디지털	플랫폼	경영	인공지능
접근	투자	미래	윤리
운영	스타트업	이사회	펀드
플랫폼	매출	책임	공동
지원	지원	글로벌	알고리즘
데이터센터	계획	역할	지원
개선	모빌리티	성장	경영
참여	협력	데이터센터	ICT
교육	재생에너지	공동체	협력
전통시장	성장	사업	계획
정보	콘텐츠	선임	투자
소상공인	전기차	공동	혁신
전환	혁신	강화	채권
계획	일자리	지속가능	조정
온라인	청년	고민	활동
노력	공동	신뢰	사업
경영	상생	평가	발전
채널	교육	논란	교육
이용자	성출	혁신	로봇
보고서	RE100	위원회	공유

<표 13> IT 분야 토픽 모델링 결과

05 결론

◆ 연구 결과 정리

- 연구문제 1 : ESG 보고서와 뉴스 기사 텍스트 분석을 통해 도출된 산업별 ESG 경영 전략 및 ESG 요소별 핵심 이슈는 무엇인가?

1) ESG 보고서

- 금융 분야: 고객 가치 제고를 최우선 가치로 삼아 성장 방향성 수립, 기후변화 대응 전략 강화
- 제조 분야: 구성원의 안전·보건을 중요시하고, 지속가능한 공급망 실현에 힘씀
- IT 분야: 기술을 통해 사회적 가치 제고
- 환경(E): 탄소중립 노력, 친환경 활동 추진 / 사회(S): 이해관계자 중심 경영 / 지배구조(G): 투명한 지배구조, 리스크 관리

2) 뉴스 기사

- 금융 분야: 환경 분야 사업 투자 확대, 친환경 금융 추진, 기업 내외로 포용 금융 실천
- 제조 분야: 저탄소 기술 개발·친환경 투자 강화, 공급망 내 ESG 관리를 통해 지속가능한 공급망 실현 노력
- IT 분야: 데이터 센터의 탄소 배출 저감 통한 탄소중립 노력, 기술 윤리·디지털 책임 강조
- 환경(E): 그린 파트너십, 탄소저감 노력 / 사회(S): 동반성장·스타트업 육성, 품질·서비스 관리 / 지배구조(G): 이해관계자 소통

17

05 결론

◆ 연구 결과 정리

- 연구문제 2 : ESG 보고서를 통해 공시하는 경영 전략과 뉴스 기사에 나타나는 ESG 이슈는 어떠한 관계성이 있는가?

- 기업에서 ESG 보고서를 통해 공시하는 지속가능경영 전략과 뉴스 기사에서 나타난 ESG 관련 이슈는 대체로 일맥상통하는 경향이 있음

보고서: 금융 분야에서 고객 중심의 전략 방향성, 친환경 금융 강조

뉴스 기사: 금융 분야 토픽 모델링 결과 '토픽 4'는 '고객 참여를 통한 친환경 사회공헌활동'으로 나타남

"KB증권, 여름철 바다 지키기 플로깅 이벤트 실시," 『EBN 산업경제』, 2022.07.04
 "신한카드, 소비자 탄소중립 참여 확대 위한 MOU," 『신아일보』, 2022.12.01

- 기업 이미지에 부정적으로 작용하는 사건에 대한 언론 보도 이후 해당 문제와 관련된 ESG 활동을 강화하는 경향이 드러남

IT 분야 기업인 네이버에서 '직장내 괴롭힘' 문제가 불거짐에 따라 조직문화와 노동 감수성에 대한 지적이 나타남

↓
 보고서: '2021 네이버 통합보고서'에서 '거버넌스 영역'의 이슈 중요도 상승
 뉴스 기사: C레벨 임원(CEO, CFO, CCO, COO)체계 전면 개편, 사건 책임자 사퇴

18

05 결론

◆ 연구 결과 정리

- 연구문제 3 : 산업별 ESG 경영 전략과 트렌드의 종합적 탐구를 통해 얻을 수 있는 전략적 시사점은 무엇인가?
 - 1) 금융 분야
 - ESG 활동을 유인할 수 있는 경영 전략을 수립할 필요가 있음
 - 지속 가능한 ESG 자원 배분을 목표로 공정한 ESG 활동 평가 체계와 인프라 구축되어야 함
 - 2) 제조 분야
 - 모든 이해관계자의 요구를 충족시키는 지속 가능한 공급망을 마련하는 데 최선을 다해야 함
 - 저탄소 전략 수립·실행과 친환경 투자를 통해 탄소 중립 활동을 지속해야 함
 - 3) IT 분야
 - 디지털 책임에 대한 소명을 갖고 건강한 디지털 문화 확산에 기여해야 함
 - 모든 이용자가 동등한 서비스를 제공받을 수 있도록 서비스 접근성을 제고해야 함

◆ ◆ ◆
감사합니다

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 ICT혁신인재 4.0 사업의 연구결과로 수행되었음”
 (IITP-2023-RS-2022-00156299)

사이버 위협 대응 모의훈련 통합형 평가체계 개발

박진용, 이용수, 김태성(충북대학교)

※ 본 논문은 저자의 요청으로 공개하지 않음.

EU 공급망 실사법 대응을 위한 중소기업 ESG 평가지표 개선방안 연구

2023년 제25회 융합학술대회 한국경영정보학회 & 한국생산관리학회

숙명여자대학교 기후환경융합학과 김수지

숙명여자대학교 산학협력단 기준학

* 본 연구는 환경부 재원으로 환경산업기술원 지식기반 환경서비스 특성화대학원 사업의 지원에 의해 수행되었습니다



목차

1. 연구의 개요
2. 연구 방법
3. 이론적 배경 및 선행연구 조사
4. 연구 수행
5. 연구 결과
6. 연구의 시사점 및 한계

연구의 필요성 및 목적

ESG경영의 부상

- 기업의 영리 활동으로 환경 악화와 기후변화의 이상 징후가 지속되면서 기업은 재무성과를 위한 활동을 넘어 환경 보호와 기후변화에 대한 책임을 수반하게 되었고, 비재무적 성과지표가 투자자들 사이에서 중요한 가치가 되었음

중소기업의 ESG 평가

- 지속 가능 금융 시장이 성장하면서 ESG 평가체계는 ESG 경영을 위한 전문인력 및 설비 투자가 가능한 대기업에는 기회 요인으로 활용될 수 있으나, 중소기업은 ESG 관련 데이터가 불충분하다는 이유만으로 투자 대상에서 제외되는 경우가 발생함

중소기업의 ESG 경영 및 평가지침의 필요성

- 대한상공회의소에서 시행한 '공급망 ESG 대응현황 조사보고서'에 따르면 국내 수출 중소기업 52.2%가 향후 공급망 내 ESG 경영 수준 미흡으로 고객사로부터 계약 수주가 파기될 가능성이 높다고 답하였으며, 수출중소기업의 ESG경영은 기초적인 수준에 머물러 있음
- 현재까지 ESG 경영의 필요성을 인식시키고, 주요 ESG 관리 항목을 제시하는 수준 ESG 항목을 성과 점검 지표와 세부 운영 방법을 제시한 지침서는 산업부의 K-ESG 가이드라인 v.1으로 한정되며, 공급망 실사법에 대비하기 위한 지침으로는 공급망 대응 K-ESG 가이드라인이 유일함

국내 ESG 정보공시 가이드라인

K-ESG 가이드라인 v.1 (산업부) 중소/중견 진단항목	ESG 벤처투자 표준지침 (중소기업벤처기업부)
공급망 대응 K-ESG 가이드라인 (산업부)	ESG 자가진단체크리스트 (중소기업ESG추진방안, 관계부처)
ESG정보공개가이드(한국거래소)	중소기업 ESG 관리 주요지표 (KPMG, 상공회의소)

1) <https://www.esgeconomy.com/news/articleView.html?idxno=2979>
 2) <https://www.kbiz.or.kr/ko/contents/bbs/view.do?seq=153559&mnSeq=207&schFld=whle&schTxt=ESG>



연구의 필요성 및 목적

EU 공급망 실사대응을 위한 중소기업의 ESG 경영 및 평가지표 제시

- 하지만 K-ESG가이드라인의 성과 지표를 중소기업에 그대로 적용하여 관리하기에는 지표 수가 많고 높은 관리 수준을 요구하고 있어 중소기업의 규모, 조직, 지배구조의 취약성에 따른 현실적인 어려움이 예상됨
- 본 연구에서는 공급망 실사 법제화가 가장 적극적으로 진행되고 있는 EU 기업 지속가능성 실사 지침 (CSDD, 이하 EU 공급망 실사법)* 대응 관점과 중소기업의 규모적 특성을 반영하여 ESG 평가지표를 제시 하고, 해당 지표를 사용해 기업을 평가하여 중소기업형 ESG 평가항목의 타당성을 추가로 분석함

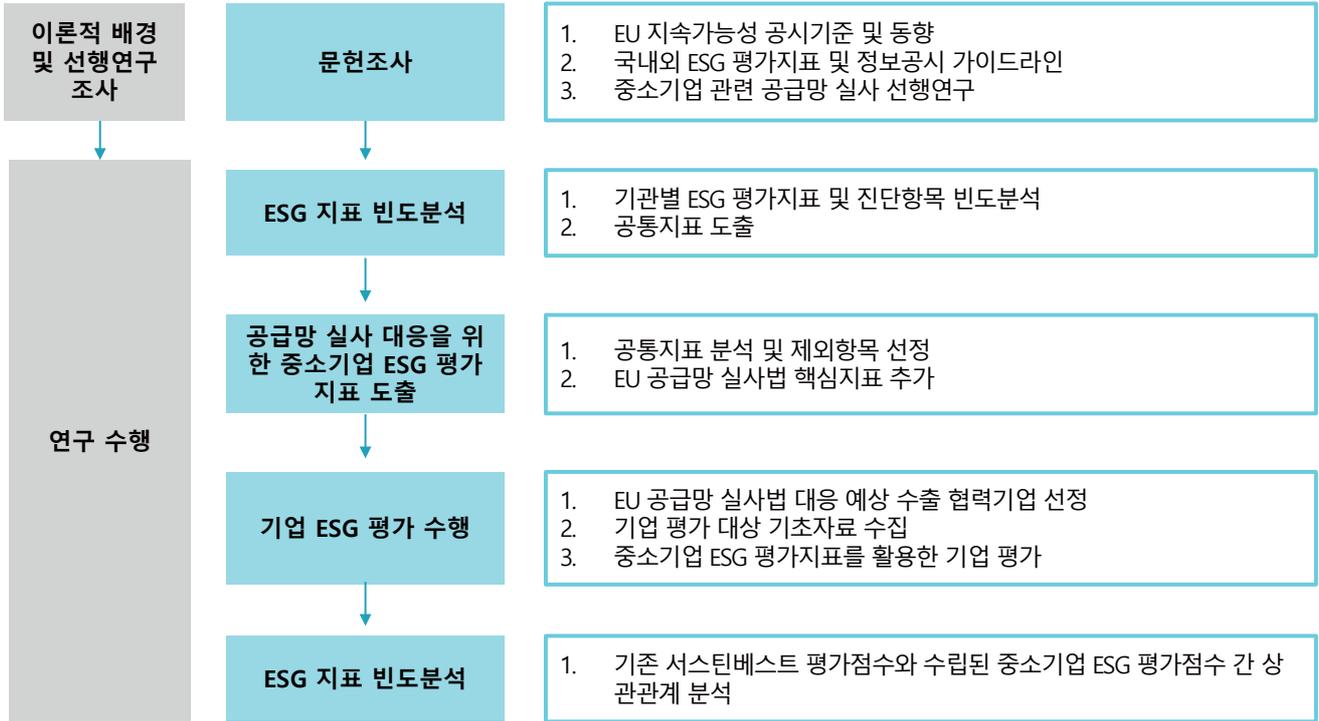
*EU 지속가능성 기업실사지침(CSDD)

'22년 2월 EU 집행위원회 채택
'23년 6월 EU 의회 표결 통과

- EU집행위원회(EU Commission)은 '22년 2월 글로벌 기업들이 공동의 기준에 맞게 경영활동을 수행할 수 있도록 지속가능한 기업실사지침 (Directive on Corporate Sustainability Due Diligence)을 채택
- 기업 공급망 내에서 발생할 수 있는 인권/환경/사회 부문의 부정적인 영향을 파악하고, 예방에 목적을 두는 ESG 공급망 실사법안이 글로벌 ESG 이슈로 떠오르고 있음
- 국내 대기업은 공급망 ESG 관리를 강화하는 추세이며, 중소기업중앙회 '23년 1월 보도자료에 따르면 주요 대기업 30개사 중 26개 기업이 협력사에 대한 ESG 평가를 실시하고 있음 (19년 17개사(56.7%) → 20년 20개사(66.7%) → 21년 26개사(86.7%)²⁾



연구 방법



문헌조사

1. EU 지속가능성 공시기준 및 동향

EU 기업 지속가능성 실사지침 (CSDD)¹⁾

Corporate Sustainability Due
Diligence Directive

- 2022년 2월, EU집행위(EC)는 공급망상 인권·환경 보호 강화를 위한 '기업 지속가능성 실사 지침' 초안을 발표하였으며, EU 회원국은 지침 발효 후 2년 이내 국내 법으로 전환, 적용해야 함
- 주요 이슈
 - 인권 : 강제노동, 아동노동, 부적절한 작업장 건강 및 안전, 노동자 착취
 - 환경 : 온실가스 배출, 오염 또는 생물다양성 손실 및 생태계 파괴, 기후변화 대응
- 적용대상 기업
 - EU Group 1 : 500인 이상 기업, 또는 매출 1.5억 유로 이상 대기업
 - EU Group 2 : 250인 이상, 매출 4000만 유로 이상 기업, 고위험업종 비중 50% 이상 중견기업
 - Non-EU : 매출 1.5억 유로 이상 대기업, 매출 4000만 초과 고위험업종 중견기업
- 실사 내용 : **기업 정책에 공급망실사 의무 통합**, 인권 및 환경에 대한 부정적 영향 파악 및 평가, 실질적/잠재적 영향의 예방/제거/최소화, 피해 구제 절차 마련, 정책 효과성 모니터링, 실사 결과에 대한 외부 공개
- EU 회원국이 지정하는 관계 당국으로부터 지침 위반시 벌금 등 행정제재를 받을 수 있음



<그림> EU 지속가능성 규제의 구조

기업책임경영을 위한 OECD 기업실사 지침 (2018) : 기업실사에 대한 권고 및 가이드라인, 기업들은 기업운영, 공급망 및 기타 사업관계와 연계되어 있는 근로자, 인권, 환경, 뇌물수수, 소비자 및 기업지배구조 등에서 발생할 부정적 영향을 방지하고 해결할 수 있도록 함
기업과 인권에 대한 이행 원칙 UNGP (2011) : 기업으로부터 개인의 인권을 보장하기 위해 국가와 기업에 인권 존중과 보호 그리고 구제에 관한 의무와 책임을 규정한 기본원칙

1) https://commission.europa.eu/publications/proposal-directive-corporate-sustainability-due-diligence-and-annex_en
2) UNGP : 인권경영이행원칙, OECD 기업책임경영 가이드라인



문헌조사

1. EU 지속가능성 공시기준 및 동향

EU 지속가능성 보고 지침(CSRD) Corporate Sustainability Reporting Directive)

- EU는 2017년부터 기업의 비재무정보공개지침인 NFRD(Non-Financial Reporting Directive)를 시행 해왔으며, 2050년 탄소중립달성을 위한 EU그린딜이 2019년에 마련된 이래, 비재무정보공시의 중요성과 이해관계자의 요구가 증대되면서, 표준화된 공시 기준 필요성이 증가함
- '21년 4월, NFRD의 개정판인 EU 지속가능성 보고지침(CSRD) 제안서가 공개되었으며, 22년 11월 EU 집행위원회로부터 최종 승인됨
- 유럽재무보고자문그룹(EFRAG)은 CSRD의 지속가능성 보고 표준안으로, '22년 11월, CSRD 이행을 위한 관리 도구로 활용될 수 있는 EU지속가능성공시표준(ESRS) 최종안을 공개함



<그림> EU 지속가능성 규제의 구조



*NFRD 공시 사항 : 환경 문제, 사회적 문제 및 직원 처우, 인권 존중, 반부패 및 뇌물수수, 회사 이사회의 다양성(연령, 성별, 교육 및 직업적 배경 측면에서)
 *CSRD는 Regulation(법률)이 아닌 Directive(지침) 형식의 입법으로 EU 회원국에 대해 직접적인 구속력은 갖지 않으며, 각 회원국은 18개월 이내에 자국 법률을 제정할 의무를 부담



문헌조사

1. EU 지속가능성 공시기준 및 동향

EU 지속가능성 보고 표준 (ESRS)

*유럽재무보고자문그룹 European Financial Reporting Advisory Group

- CSRD 대상 기업은 EU지속가능성공시표준(ESRS)에 따라 보고해야 함
- ESRS는 공통기준 2가지 ESRS1(일반요구사항), ESRS2(일반, 전략, 거버넌스 및 중요성 평가 공시 요구사항)로 구성되어 있으며 10가지 ESG 주제별 기준을 포함하여 기업이 공시할 정보의 범위와 기준을 명시하고 있음



<그림> EU 지속가능성 규제의 구조



ESRS 2에 명시된 중요성 평가 기준에 따라 해당 토픽의 중요성을 평가하고 중요한 항목에 대해 우측 보고 영역을 공시

<그림> ESRS의 공시 항목의 기본 구성 (자료: pwc)



문헌조사

2. 국내외 ESG 평가지표 및 정보공시 가이드라인

ESG 관련 지표 조사



- 국내외 ESG 평가지표 및 정보공시 가이드라인 문헌조사 수행
- 중소기업에 대상으로 공개된 ESG 정보공시 가이드라인은 총 6개로 확인하였으며, 이 중 산업부의 공급망대응 K-ESG 가이드라인은 중견·중소 협력기업의 공급망 실사 대응을 위해 제작됨
- 산업부는 중소·중견기업의 현장의견을 수렴하여 이를 바탕으로 공급망 대응 K-ESG 가이드라인 진단항목정의서 고도화 과정을 이행하였으며, 진단 체계를 기초(3개 영역, 25개 범주, 25개 항목)와 심화 항목(4개 영역, 18개 범주, 60개 항목)으로 이원화 하였음

<표> 국내외 ESG 평가기관 및 정보공시 가이드라인

국내 평가기관	정보공시 가이드라인	해외 평가기관
한국ESG기준원(KCGS)	K-ESG 가이드라인 v.1 (산업통상자원부)	MSCI 모건스탠리
서스틴베스트	공급망 대응 K-ESG 가이드라인 (산업통상자원부)	DJSI S&P 다우존스
한경 ESG 경영대상 평가지표	ESG정보공개가이드선(한국거래소)	Moody's 무디스
	ESG 벤처투자 표준지침 (중소기업벤처기업부)	Refinitiv 레피니티브
	ESG 자가진단체크리스트 (중소기업ESG촉진방안, 관계부처)	Reprisk 랩리스크 ESG ratings
	중소기업 ESG 관리 주요지표 (KPMG, 대한상공회의소)	EcoVadis
	3-Step Sustainability Assessment for SMEs(EU 회계사협회)	



Step 1. 기존 ESG 지표 빈도분석

공통지표 추출

- ESG 평가지표 및 ESG 정보공시 가이드라인 진단항목을 총 조사하여 빈도분석 수행
- 분석대상 : 총 16개 전문기관의 ESG 지표, 5개 이상 (30%) 공통지표 추출 → **중요성 확인**

분류	관리 Index	합계	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
ESG총괄	지속가능경영 관련 인증	8																
	ESG 정보공시 방식	5																
	온실가스 배출량(Scope 1, 2)	14																
	온실가스 배출량(Scope 3)	11																
정보공시	용수 사용량	11																
	환경 에너지 사용량	10																
	온실가스 배출량 검증	9																
	폐기물 배출량	9																
	폐기물 재활용 비율	9																
	대기오염물질 배출량	6																
	환경 법/규제 위반	6																
	생물다양성 & 토지사용	6																
	수질오염물질 배출량	5																
	원부자재 사용량	5																
	온실가스 배출 감축 노력 (에너지 사용 효율화 등)	5																
	재생에너지 사용 비율	5																
	유해물질 배출/폐기물 최소화	5																
친환경 인증 제품 및 서비스 비율	5																	
제품 탄소발자국 탄소배출량 측정 및 관리	5																	
정보보호 시스템 구축	10																	
근로자의 보건안전 화재 대피 훈련, 업무중 휴식 장려 등	9																	
사회	여성 구성원 비율	8																
	장애인 고용률	7																
	구성원 봉사참여	7																
	개인정보 침해 및 규제	7																
	안전사고 근절을 위한 품질검사, 대외 품질인증 획득	7																
	자사/공급망 미성년 근로자 부적법 채용, 취약계층 강제노동	7																
	적절한 근로조건	7																
	산업재해율	6																
	인관정책 수립	6																
	이해관계인 등과 부당하거나 불공정한 거래	6																
	고용상의 차별 금지	5																
	근로계약서 작성 및 최저임금 지불	5																
	지역공급자 또는 지역 인재 배려 정책 창업 지원	5																
정규직 비율	5																	
지배구조	비율리 행위 예방 조치(부패, 뇌물수수, 부정청탁)	8																
	사외이사 비율	6																
	대표이사 이사회 의장 분리	6																
	윤리규범 위반사항 공시	5																
	윤리현장 및 실천규범	5																
이사회 다양성	5																	



* A~P까지 순서대로 : Refinitive, K-ESG, 공급망 대응K-ESG, KCGS, 벤처투자ESG, S&P, Moody's, 회계사협회, Reprisk, 서스틴베스트, MSCI, 대한상공회의소, 한국거래소, ESG자가진단체크리스트, 한경 경영대상, EcoVadis



Step 2. 중소기업형 ESG 평가지표 선정

중소기업형 ESG 지표 선정

- 중소기업은 대기업과 달리 규모 특성상 원활한 ESG 경영의 도입에 내부적인 제약이 존재함
- 기업 경영인이 기업의 소유자와 일치, 비상장기업으로 대리인 문제 없음, 주주권 보호와 배당 확대를 위한 경영상 장치 불필요 (지배구조 특성)
- ESG 전문성 확보 및 체계적인 역할분배 어려움(규모적 특성)

영역	평가 항목	구분
ESG총괄	지속가능경영 관련 인증	○
정보공시	ESG 정보공시 방식	○
환경	온실가스 배출량 Scope 1,2	○
	온실가스 배출량 Scope 3	△
	용수 사용량	○
	에너지 사용량	○
	온실가스 배출량 검증	○
	폐기물 배출량	○
	폐기물 재활용 비율	○
	대기오염물질 배출량	○
	환경 법/규제 위반	○
	생물다양성&토지사용	△
	수질오염물질 배출량	○
	원부자재 사용량	○
	온실가스 배출 감축 노력	▲
	재생에너지 사용 비율	▲
	유해물질 배출 및 폐기물 최소화	○
	친환경 인증 제품 및 서비스 비율	○
	제품 탄소발자국 탄소배출량 측정 및 관리	△

영역	평가 항목	구분
사회	정보보호 시스템 구축	○
	근로자 보건안전 화재 대피 훈련 업무중 휴식 장려 등	○
	여성 구성원 비율	○
	장애인 고용률	▲
	구성원 봉사 참여	♣
	개인정보 침해 및 규제	○
	안전사고 근절을 위한 품질검사, 대외 품질인증 획득	○
	자사/공급망 미성년 근로자 부적법 채 용, 강제노동	○
	적법한 근로조건	○
	산업재해율	○
	인권정책 수립	○
	이해관계인 등과 부당하거나 불공정한 거래	○
	고용상의 차별 금지	○
	근로계약서 작성 및 최저임금 지불	○
	지역공급자 또는 지역 인재 배려 정책 창업 지원	♣
	정규직 비율	○

영역	평가 항목	구분
지배구조	비윤리 행위 예방 조치 (부패, 뇌물수수, 부정청탁)	○
	사외이사 비율	○
	대표이사 이사회 의장 분리	○
	윤리규범 위반사항 공시	○
	윤리현장 및 실천규범	○
	이사회 다양성	♣

- : 중소기업형 지표
- △ : 단계적 적용(제한된 인력 및 전문성)
- ▲ : 단계적 적용(자본과 자원의 부족)
- ♣ : 단계적 적용(업계 구조와 문화)
- ♣ : 단계적 적용(단기적 이익 추구)



Step 3. EU 공급망 실사 지표 선정

공급망 실사 핵심지표 선정

- 공통지표 + ESRS 공시항목에 따른 ESG 공급망 핵심지표 선정

※ ESRS 공시 요구사항

E1 : 기후변화, E2 : 오염, E3 : 물과 해양자원, E4 : 생물다양성, E5 : 자원사용과 순환경제
S1 : 인력, S2 : 가치사슬 내 근로자(중소기업의 경우 해당X), S3 : 지역사회, S4 : 소비자, G1 : 사업수행

영역	평가 항목	분류
ESG총괄	지속가능경영 관련 인증	반도분석 최다지표
정보공시	ESG 정보공시 방식	반도분석 최다지표
환경	환경 교육	중요지표 추가
	온실가스 배출량 Scope 1,2	E4
	용수 사용량	E1
	에너지 사용량	E1
	온실가스 배출량 검증	E1
	폐기물 배출량	E5
	폐기물 재활용 비율	E5
	대기오염물질 배출량	E2
	환경 법/규제 위반	E1
	수질오염물질 배출량	E2
	원부자재 사용량	E5
	유해물질 배출 및 폐기물 최소화	E2
	친환경 인증 제품 및 서비스 비율	E5

영역	평가 항목	분류
사회	인권정책	S1
	인권교육	S1
	결사의 자유 보장	S1
	아동노동/강제노동 금지	S1
	근무시간 준수	S1
	여성 임직원 비율	S1
	산업안전보건체계	S1
	산업재해율	S1
	작업환경	S1
	품질 인증	중요지표 추가
	개인정보 침해 및 규제	S1
	지역사회 공헌	S3

영역	평가 항목	분류
지배구조	사외이사 비율	중요지표 추가
	경영정보공시	G1
	ESG 정보공시	중요지표 추가
	윤리경영 지침 수립 여부	G1
	기업윤리교육	G1
	감사기구 구축	G1
	내부회계관리제도	중요지표 추가



Step 4. ESG 지표 배점 설정

ESG 지표별

평가 배점 분배

- 환경, 사회, 지배구조 부문의 만점은 100점, 총 점수 300점 (부문별 위험요소 -10점)
- 서울연구원, AHP 조사결과 「서울시 소재 소재기업 ESG 관련 정책수요조사 보고서」에서 '수출기업'의 ESG 부문별 중요도 순위를 참고하여 배점 분배



E(환경) 세부지표별 순위					
환경경영	온실가스	에너지	용수	폐기물	오염물질
39.3	21.3	13.1	4.9	4.9	16.4
S(사회) 세부지표별 순위					
정규직비율	결사의 자유 보장	여성 구성원 비율	장애인 고용률	산업재해율	전략적 사회 공헌
21.3	41	26.2	6.6	1.6	3.3
G(지배구조) 세부지표별 순위					
이사회구성	이사회활동	주주관리	윤리경영	감사기구	지배구조 법/규제 위반
1.6	11.5	23	29.5	24.6	9.8

+

- E1 : 기후변화
 E2 : 오염
 E3 : 물과 해양자원
 E4 : 생물다양성
 E5 : 자원사용과 순환경제
- S1 : 인력
 S2 : 가치사슬 내 근로자
 S3 : 지역사회
 S4 : 소비자
- G1 : 사업수행

자료: 서울연구원

수출기업 ESG 중요도 순위(%)

ESRS 공시 요구사항
(공급망 실사 핵심 지표 추가)



Step 4. ESG 지표 배점 설정

ESG 지표별

평가 배점 분배

- 서울연구원 수출기업 ESG 중요도를 참고하여 환경, 사회, 지배구조 세부배점을 분배하였으며, ESRS의 핵심항목을 추가하여 배점을 부여함
- 수집된 정보에 따라 ESG 평가점수를 5단계(매우 높음, 높음, 보통, 낮음, 매우 낮음)으로 평가를 시행하였음

E(환경) 부문 지표 배점 조정					S(사회) 부문 지표 배점 조정					G(지배구조) 부문 지표 배점 조정				
대분류	서울연구원 중요도 순위	중소기업 평가배점	ESRS 표준	보완	대분류	서울연구원 중요도 순위	중소기업 평가배점	ESRS 표준	보완	대분류	서울연구원 중요도 순위	중소기업 평가배점	ESRS 표준	보완
환경경영	39.3	35	E0 총괄		인권경영	95.1	55	S1 인력		투명경영	-	35	G1 사업수행	+ 실사
온실가스	21.3	20	E1 기후변화		인권보호	41.0	30	S1 인력		윤리경영	29.5	35	G1 사업수행	
에너지	13.1	15	E1 기후변화		인적자원	21.3	15	S1 인력		윤리경영	29.5	35	G1 사업수행	
원자재	-	5	E5 자원순환	+ 실사	다양성 및 양성평등	32.8	10	S1 인력		감사제도	24.6	20		
생물다양성	-	5	E4 생물다양성	+ 실사	안전보건	1.6	30	S1 인력		이사회 구성	1.6	10		
용수	4.9	5	E3 수자원		제품품질	-	5	S4 소비자 및 최종사용자	+ 실사	이사회 활동	11.5	-		-제외
폐기물	4.9	5	E5 자원순환		정보보호	-	5	S1 인력, S4 소비자 및 최종사용자	+ 실사	주주관리	23.0	-		-제외
오염물질	16.4	10	E2 오염		지역사회	3.3	5	S3 영향권 내 지역사회		지배구조 법/규제위반	9.8	-		별도 구성
합계	100	100			합계	100	100			합계	100	100		

인권보호 이슈 모두를 대외공식적 정책으로 제시하기에는 한계가 있으므로, 인권경영 하위, 핵심지표로서 인권보호, 인적자원, 다양성 등의 성과가 있으면 점수를 받을 수 있도록 구성



Step 5. 중소기업 ESG 평가

기업 ESG 평가 실행

- 목적 : EU 공급망실사 대응을 위한 중소기업 ESG 평가지표의 타당성 확인
- 평가 대상 : 서스틴베스트 ESGValue 2021년 기업평가 대상 기업(반도체, 자동차, 석유, 제약 및 바이오)
- 결과 분석 : 서스틴베스트 ESGValue 점수와 부문별 상관관계 분석

평가 기업 선정

서스틴베스트 ESGValue 평가 자료(2021)



- ① 서스틴베스트 ESGValue 2021년 기업평가자료 기반
- ② 중소기업(10개사), 중견기업(7개사)
- ③ ESG 비재무 정보를 공개하고 있는 기업을 우선순위로 선정(보고서, 홈페이지, 미디어 등)

기업 평가

EU 공급망실사 대응을 위한 중소기업 ESG 평가지표



- ① 17개사 대상 기초자료 수집 (보고서, IR 기술분석보고서, 홈페이지, 미디어 뉴스)
- ② 수집된 정보의 수준에 따라 배점 부여 (매우 높음, 높음, 보통, 낮음, 매우 낮음)
- ③ 공개되지 않은 정보에 대해서는 0점 처리

평가 결과 분석

중소기업 ESG 평가지표 vs. ESGValue 평가자료



- ① 평가지표 배점 및 활용 타당성을 파악하기 위해 서스틴베스트 점수와 중소기업 평가지표 원점수 비교



중소기업 ESG 평가 결과

부문별 평가 결과

- 본 연구의 환경, 사회, 지배구조 부문 평균 점수는 서스틴베스트 결과에 비해 각각 273%, 52%, 2% 상승하였고, ESG 평가 원점수 총점 평균 점수는 145점으로, 서스틴베스트 108점보다 35% 상승함
- 전체적으로 본 연구에서 수립한 평가지표를 적용한 결과가 서스틴베스트에 비해 높게 나타난 것은 평가지표 구성에 따라 중소기업이 상대적인 차별을 받을 가능성을 보여주는 결과이며, 연구의 타당성을 확인함

환경	본연구(A)	서스틴베스트(B)	증감률(A/B*100)
1	69	39	75%
2	14	6	120%
3	89	71	26%
4	74	55	36%
5	40	16	158%
6	83	54	53%
7	40	18	129%
8	38	28	36%
9	2	1	100%
10	44	20	116%
11	22	10	129%
12	41	3	1,104%
13	33	3	876%
14	32.7	4.5	627%
15	37.5	5.29	609%
16	46	14	216%
17	36	11	223%
평균	43	21	273%

사회	본연구(A)	서스틴베스트(B)	증감률(A/B*100)
1	41	21	100%
2	43	41	6%
3	51	46	12%
4	48	52	-9%
5	32	30	7%
6	101	47	114%
7	15	25	-40%
8	91	44	105%
9	35	10	233%
10	30	16	91%
11	38	28	34%
12	43	19	121%
13	85	60	41%
14	25	14.29	75%
15	22.5	31.8	-29%
16	48	50	-4%
17	55	44	24%
평균	47	34	52%

지배구조	본연구(A)	서스틴베스트(B)	증감률(A/B*100)
1	63	49	30%
2	57	58	-2%
3	44	50	-11%
4	79	58	36%
5	64	60	6%
6	57	44	29%
7	55	48	13%
8	62	58	6%
9	45	58	-23%
10	38	45	-16%
11	65	59	10%
12	64	52	22%
13	51	45	12%
14	19.5	44.07	-56%
15	52.5	53.64	-2%
16	62	63	-2%
17	49	58	-16%
평균	54	53	2%

▶ 중소기업의 경우 대기업보다 생산공정에 집중하고 있으며, 중소기업 ESG 평가지표는 서스틴베스트와 비교하여 생산공정에 높은 배점을 부여하고 있으므로 중소기업의 ESG 평가점수가 높게 도출된 것은 중소기업형 지표로서 타당성을 가지는 것으로 볼 수 있음

▶ 서스틴베스트의 사회 평가항목 중 국제 이니셔티브 가입 및 활동, 공정거래, 상생협력, 공급사슬관리 평가항목은 주로 대기업에 적용되는 평가지표로서, 중소기업 ESG 평가점수가 기존 대비 높게 도출된 것은 중소기업의 특성과 역량을 고려한 결과로 볼 수 있음

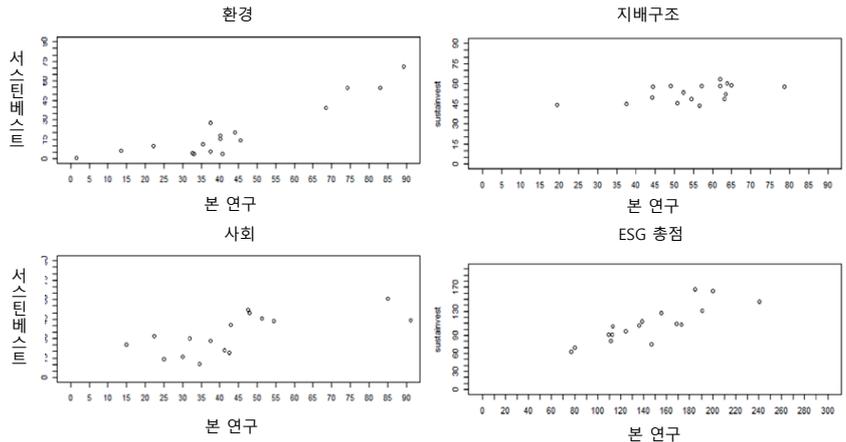
▶ 주주 권리, 이사회 세부 활동 등 항목이 제외되어 서스틴베스트 평가 대비 낮은 점수를 기대하였으나, 중소기업 특성을 고려하여 상기 항목에 대해 가중치를 작게 반영하여 점수 차이가 크지 않은 것으로 분석됨

중소기업 ESG 평가 결과

상관관계 분석

- R을 활용한 서스틴베스트 평가 점수와 ESG 중소기업 평가 점수의 상관관계 분석 시행
- 환경(E) 부문의 두 변수 간의 Pearson 상관관계를 분석한 결과는 0.9068이며, 이 값은 두 변수 간 강한 선형 관계가 있으며, p-value는 5.193e-07 < 0.05 값을 만족하므로 변수간 상관관계가 통계적으로 유의미한 결과
- 사회(S) 부문의 두 변수 간의 Pearson 상관관계를 분석한 결과는 0.6756이며, 이 값은 두 변수 간 선형 관계가 있으며, p-value는 2.915e-03 < 0.05 값을 만족하므로 변수간 상관관계가 통계적으로 유의미한 결과
- 지배구조(G) 부문의 두 변수 간의 Pearson 상관관계를 분석한 결과는 0.5430이며, 이 값은 두 변수 간 선형 관계가 있으며, p-value는 0.0243 < 0.05 값을 만족하므로 변수간 상관관계가 통계적으로 유의미한 결과

영역	상관계수	t-값	p-value
환경(E)	0.9068	8.333	5.193e-7
사회(S)	0.6756	3.549	2.915e-3
지배구조(G)	0.5430	2.504	0.02430
ESG 총점	0.8414	6.031	2.301e-5



중소기업 ESG 평가 결과

ESG 평가 세부항목

점수등급별 분석

- 중소중견 기업 17개사를 대상으로 세부 항목별 백분위 점수 평균값을 산정하여 ESG 항목별 중소기업 대응 필요성을 위험도 등급으로 표현함
- 위험도 등급 : D 매우 높음(0~10점), C 높음(10~30점), B 보통(30~60점), A 낮음(60~90점), S 매우 낮음(90점 이상)
- D등급의 경우, 평가항목 대응 기반이 구축되지 않았거나 인력 및 자본이 부족하므로, 장기적인 대응 전략필요
- C등급의 경우, 활동을 수행하지만, 정량적인 모니터링을 하지 않거나, 국내 법규 대응에만 주력하여 규제를 회피하는 방식으로 나타나, 공급망 대응을 위해 국내 법규 뿐 아니라 글로벌 원청기업의 요구에 따라 대응필요
- A, B등급의 경우 리스크 대응 방향을 유지하고 지속적인 모니터링 및 필요한 부분을 개선하여야 함

점수	등급	위험도	평가항목
90~	S	매우 낮음 (Very Low)	내부회계관리제도, 정규직비율, 경영정보 공시
60~90	A	낮음 (Low)	여성 임직원 비율, 산업안전보건 체계, 지역사회 공헌, 윤리경영 지침 수립 여부, 환경경영시스템, 품질 인증
30~60	B	보통 (Moderate)	에너지 사용량, 폐기물 배출량 관리, 기업윤리교육, 환경교육, 감사기구 구축, 사외이사 비율, 개인정보 침해 및 구제
10~30	C	높음 (High)	폐기물 재활용, 폐수 처리량 관리, 원부자재 사용량, 산업재해율, 근무시간 준수, 수질오염물질 배출량, 작업환경, 대기오염물질 배출량, 온실가스 배출량(Scope 1, 2), ESG정보 공시, 용수 사용량 관리
0~10	D	매우 높음 (Very High)	아동노동/강제노동 금지, 윤리규범 위반사항 공시, 인권교육, 결사의 자유 보장, 생물다양성, 인권정책

* 위험도가 높은 C, D등급은 17개로, 전체 지표 수(33개) 중 약 51%를 차지하였음.

* 위 결과는 중소기업의 공급망 실사법 대응의 취약성을 보여주며, 평가항목에 대한 관리 방안 및 장기적인 전략 구축의 필요성이 있음



연구의 시사점 및 한계

연구결과 및 시사점

1. 국내외 ESG 평가지표 및 정보공시표준 항목을 조사하여 지표별 빈도분석을 수행하였고 이를 통해 각 분야에서 중요도가 높은 공통지표를 도출하였으며, 중소기업 특성에 따라 ESG 경영 내부 제han적 요소를 고려하여 세부 평가항목을 보완하였음
2. 개선된 중소기업 ESG 평가지표의 타당성을 확인하기 위하여 중소중견 기업을 대상으로 ESG 평가를 시행 하였음¹⁾
3. 원점수 기준으로 중소기업 ESG 평가결과와 서스틴베스트 ESGValue 평가결과를 분석한 결과, 중소기업 평가지표를 적용할 때, 서스틴베스트 평가 원점수에 비해 평균 35% 상승한 것을 확인함
이러한 결과는 중소기업이 ESG 등급 평가에서 대기업과 중견기업에 비해 불리한 위치에 있는 것을 보여 주며, 중소기업이 고유한 규모와 특성을 고려한 평가체계가 필요함을 시사함
4. 본 연구는 중소기업의 규모적 특성과 현실 여건을 반영하여 현장 적용성을 높이고 과소평가를 방지하는 중소기업형 ESG 지표를 제시하였다는 점과 중소기업에 적합한 EU 공급망 실사법 요구사항을 적용하여 ESG 평가지표를 수립함으로써 중소기업의 대응 기반을 마련하였다는 점에서 의의가 있음

연구의 한계

1. EU 공급망 실사법 대응 관점에서 중소기업의 ESG 평가지표를 사용하여 17개 기업을 평가한 결과와 서스틴베스트를 비교하였으나 대상이 제한적이었으므로 일반화의 오류가 발생할 수 있다고 판단하였음 → 더 많은 데이터(30개 이상)를 기반으로 추가 연구가 필요함
2. 기후변화 대응 중요성이 커지면서 Scope 3 배출량 공시요구가 증가하고 있으나 본 연구의 중소기업 ESG 평가지표에는 포함되지 않음 → 지속가능 ESG 경영 체제 도입 및 리스크 관리를 위해 적극 대응 필요

1) 부문별 평가지표의 배정은 서울연구원에서 수행한 「서울시 소재 기업 ESG 관련 정책수요조사 보고서」 중 '수출 중소기업을 대상으로 시행한 ESG 세부지표별 AHP 설문조사 자료를 참고함



참고문헌

김주영. (2023). 글로벌 ESG 공급망 실사 확대에 따른 협력 중소기업 ESG 지원방안 연구 - 동반성장위원회 협력사 ESG 지원사업을 중심으로 - 중앙대학교 산업창업경영대학원
김지현. (2022). EU 주요국의 공급망 ESG 실사 제도 현황 및 시사점. 월간 KIEI 산업경제, 287, 70-80.
권배성 & 이재현. (2019). 국내 중소기업의 사회적 책임 연구 동향 분석 및 향후 연구 방향 제안. 전략경영연구, 22(3), 21-50.
김성화. (2021). EU 산업정책의 전환에 따른 중소기업의 ESG 기준의 도입시 고려사항. 유통발전연구, 8(2), 203-236.
김성화. (2021). EU 산업정책의 전환에 따른 중소기업의 ESG 기준의 도입시 고려사항. 유통발전연구, 8(2), 203-236.
김효정 & 이준석. (2021). 기업의 ESG 실행과 평가의 괴리 가능성 연구: ESG 평가지표 분석을 중심으로. 국가정책연구, 35(4), 199-225.
나수미, 김주미, & 김국동. (2021). 중소기업 ESG 연구: 탄소중립을 중심으로. 중소기업정책연구보고서, 2021(31), 1-302.
대한상공회의소. 수출기업 공급망 ESG실사 대응 현황과 과제 조사(2022.07.18). http://www.korchanet.net/Cham/Service/Economy/app/KccrReportDetail.asp?SEQ_NO_C010=20120935273&CHAM_CD=8001
대한상공회의소 ESG 확산 및 정착을 위한 기업 설문조사(2021.12.29). http://www.korchanet.net/Cham/Service/Economy/app/KccrReportDetail.asp?SEQ_NO_C010=20120934572&CHAM_CD=8001
박순애 & 신은혜. (2021). 대, 중소 및 중견 기업의 ESG 성과의 재무가치 간의 상관관계 분석. 환경경제, 29(4), 151-199.
박윤수. (2023). 중소기업의 ESG 평가에 대한 전략적 대응방안 탐색적 연구. 벤처창업연구, 18(1), 47-65.
삼일회계법인(2023.01). EU CSRD 기업지속가능성보고서지침 주요 내용
서환승. (2021). 공급망 관점에서 본 수출 중소기업의 ESG 전략 도출에 관한 연구. 부산전해경제지유구역청 관내 기업을 중심으로. 동아대학교 대학원
울산산림(2011.12.29). "중소기업 장애인 모시기 비상" <https://www.usanpress.net/news/articleView.html?idxno=120499>
이연우. (2022). 한국형 ESG 가이드라인의 탐색분석과 적용포인트. Auditor Journal, 266, 2.
이인형. (2021). ESG 평가 체계 현황과 특성 분석. [KGM] 이슈보고서.
이지혜. (2022). ESG 평정의 실증분석을 통한 중견 중소기업의 ESG 경영 추진의 필요성에 대한 연구. 인문사회 21, 13(2), 2787-2802.
이지혜. (2022). ESG 평정의 실증분석을 통한 중견 중소기업의 ESG 경영 추진의 필요성에 대한 연구. 인문사회 21, 13(2), 2787-2802.
이코리야(2022.05.09) "기업 경쟁력 좌우하는 ESG경영, 중소기업 고립 왜?" <https://www.ekorenews.com/news/articleView.html?idxno=60623>
임형철 & 정우성. (2021). 국내외의 ESG 사례를 통해 본 중소기업 ESG 경영 활성화 방안. 아태비즈니스연구, 12(4), 179-192.
장해진 & 최윤정. (2021). ESG에서의 환경정보 공개의 내용 및 공시방법론에 대한 논의. 일간지, 50, 267-295.
중소기업중앙회(2022).공급망 ESG대기업협력사수출 중소기업의 대응현황 조사보고서 <https://www.kbiz.or.kr/ko/contents/bbs/view.do?mSeq=2078&seq=152394>
한국거래소(2022.04). IFRS 지속가능성 공시기준 조안 주요내용.
한국거래소(n.d). ESG 정보공개 가이드라인. <https://esg.krx.co.kr/contents/01/01030200/ESG01030200.jsp>
한국거래소(n.d). ESG의 부상배경. <https://esg.krx.co.kr/contents/01/01010200/ESG01010200.jsp>
황인창, 김고은, 백종락, 이윤혜.(2023)."환경사회지배구조(ESG)경영 확대 위한 서울시 정책방안". 서울연구원 연구보고서 2022-PR-16
"About DJSI". DJSI 홈페이지. http://djsi.org/wp/?page_id=5
Accountancy Europe(2023.03.10). 검색:"3-step sustainability assessment for SMEs" <https://www.accountancyeurope.eu/publications/3-step-sustainability-assessment-for-smes/>
EcoVadis(2023.03.10). 검색:"EcoVadis CSR Rating Methodology". https://www.ecovadis.com/theme_assets/143203/d135d5e09447424a01d1018d4411972f146dcd1.pdf
EFRAG(2022.11). "Draft European Sustainability Reporting Standards ESRS E1 Climate Change November 2022"
EFRAG(2022.11). "Draft European Sustainability Reporting Standards ESRS E2 Pollution November 2022"
EFRAG(2022.11). "Draft European Sustainability Reporting Standards ESRS E4 Biodiversity and ecosystems November 2022"
EFRAG(2022.11). "Draft European Sustainability Reporting Standards ESRS E5 Resource use and circular economy November 2022"
EFRAG(2022.11). "Draft European Sustainability Reporting Standards ESRS G1 Business Conduct"
EFRAG(2022.11). "Draft European Sustainability Reporting Standards ESRS S2 Workers in the value chain November 2022"
EFRAG(2022.11). "Draft European Sustainability Reporting Standards ESRS S3 Affected communities November 2022"
EFRAG(2022.12). "Glimpse into draft ESRS E1 Climate Change"
European Commission(2023.03.10). 검색:"Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on Corporate Sustainability Due Diligence and amending Directive (EU) 2019/1937. eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0071
European Commission(2022.12.16). "Corporate sustainability reporting Directive" https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en
Higham, I. (2019). UN guiding principles on business and human rights. In: Standardization and Risk Governance(pp. 217-234). Routledge. <https://www.msia.com/our-solutions/esg-investing/esg-ratings>
Moody's analytics(2023.3.10). 검색:"ESG Score Predictor: Applying a Quantitative Approach for Expanding Company Coverage". https://www.moodyanalytics.com/articles/pa/2022/esg_score_predictor_applying_a_quantitative_approach_for_expanding_company_coverage
MSCI(2023.3.10). 검색:"MSCI ESG Rating"
MSCI(2023.3.10). 검색:"MSCI Materiality Map" <https://www.msci.com/our-solutions/esg-investing/esg-industry-materiality-map>
OECD. (2018). OECD due diligence guidance for responsible business conduct.
Refinitiv(2023.3.10). 검색:"ESG Data" https://solutions.refinitiv.com/esg-data?utm_content=Company%20Data-KR-APAC-G-EN-BVM&utm_medium=cpc&utm_source=google&utm_campaign=596226_PaidSearchInvestmentSolutionsBAU&relqCampaignId=16987&utm_term=+esgf%20+rating&gclid=CJ0KCOAibagBhD3ARsANRqEuhSt_091VEVr0MhMmVCSuCrKWjWGF2i2H5-MZ20x0BkZG3YslaAyuEALw_wcB&gclid=cw5d
S&P Global(2023.3.10). 검색:"ESG Scores" <https://cutty/f83jnLN>
SASB. "SASB Materiality Map" <https://www.sasb.org/standards/materiality-map/>



텍스트마이닝을 활용한 기업의 ESG 중대성(Materiality)평가 정합성 분석 연구

숙명여자대학교 일반대학원
기후환경에너지학과 김해리
숙명여자대학교 기준학

본 연구는 환경부 지식기반 환경서비스 특성화대학원 사업으로 지원을 받아 수행한 과제입니다.

CONTENTS

01 연구배경 및 목적

02 이론적 배경 및 선행연구

03 연구방법

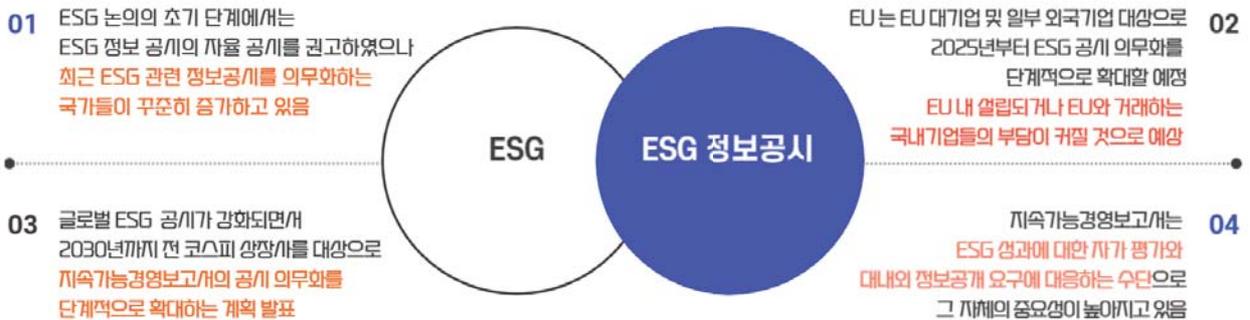
04 분석결과

01

연구 배경 및 목적

01. 연구배경 및 목적

연구배경 ESG와 ESG 정보공시

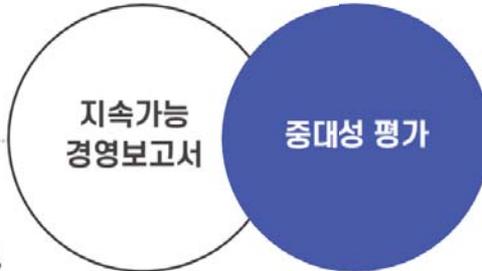


지속가능경영보고서	적용시기	1단계(~'25년)	2단계('25년~'30년)	3단계('30년~)
	자산총액	자율공시 활성화	일정규모 이상 (예: 자산 2조원 이상)	순코스피상장사

연구배경 지속가능경영보고서와 중대성 평가

01 중대성 평가는 기업의 비즈니스와 이해관계자에게 중대한 이슈를 도출하는 평가로 이를 통해 지속가능경영보고서의 공개 범위와 내용이 결정

03 국내외 ESG 가이드라인은 중대성평가를 시행하고 그 결과를 보고서 내에 공시할 것을 권고하고 있으며, 각 가이드라인별 중대성 평가 프로세스를 제시하고 있음 그러나 가이드라인별 중대성 평가 프로세스의 세부적인 방법은 상이하고 가이드라인을 반영하여 지속가능경영보고서를 작성하는 기업의 중대성평가 프로세스 또한 다양한 방식으로 진행됨



'중대성'이란 특정 이슈의 포함 여부를 결정하기 위한 기준이 되는데 중대성의 기준은 기업의 목적에 따라 상이함

04 중대성 평가 결과의 신뢰성 확보 및 ESG 워싱방지를 위해 중대성 평가 프로세스의 점검과 통일성 확보가 필요

연구 목적

- 기업의 ESG 공시와 기업의 ESG 성과 일치 여부 확인
ESG 공시 범위와 내용을 결정하는 중대성 평가 결과와 기업의 ESG 경영 성과의 정합성을 확인
- 정합성과 기업의 ESG 평가 등급과의 상관관계 확인
기업의 ESG 공시와 기업의 ESG 성과의 정합성과 기업의 ESG 평가 등급간 상관관계 확인
- 업종별 ESG 공시와 ESG 성과 특징 확인
업종별 중대 이슈 구성의 특성과 차이 확인

ESG 공시가 ESG 성과와 일치하는지 살펴보고 그 정합성이 ESG 평가등급과 어떤 상관성이 있는지 확인하여 ESG 정보 공시 발전방안에 대하여 논의하고자 함

02

이론적 배경 및 선행연구

06

02. 이론적 배경 및 선행 연구

이론적 배경 ESG와 정보공시

01

ESG는 환경, 사회, 지배구조의 약자로
지속가능성을 달성하기 위한
기업경영의 3가지 핵심요소를 의미

03

중대성 평가란 기업의 비즈니스와 이해관계자에게
중요한 중대 이슈를 도출하는 평가로
중대 이슈를 통해 지속가능경영 보고서의
공개 범위와 내용이 결정됨

02

국내 ESG 공시규제는 지속가능경영보고서 발간 기업의
연평균 20% 증가를 목표로 하고 있으며,
2025년부터는 자산총액 2조 원 이상의 코스피 상장사의
지속가능경영보고서 발간이 의무화될 예정임

04

국내외 ESG 가이드라인은 중대성 평가 실시와
결과 내용을 보고서 내에 공시할 것을 권고하고 있으며
가이드라인별 중대성 평가 프로세스를 제시하고 있음
그러나 가이드라인별, 기업별 세부 프로세스가 상이하게 이루어지므로
신뢰성 확보와 ESG 워싱 방지를 위해 중대성 평가 방식의 점검 및 통일성 확보가 필요함



07

선행연구

제목	저자(년도)	주요 내용
국민연금 기금 책임투자자와 외부 ESG 평가와의 관계에 관한 연구	정경화, 박삼복(2022)	- 지속가능보고서 공시 기업은 기업가치에 긍정적인 영향을 나타냄 - 지속가능경영을 위한 기업 비용의 일시적인 증가로 인해 단기적 수익성이 약화될 수 있으나, 기업의 자산규모에 따라 투자자들에게 ESG 정보제공을 통한 정보 비대칭성 해소와 시장에서의 신호효과로 기업의 수익성을 개선할 수 있음 - 기업의 안정성 측면에서 공시 여부에 따라 부채비율이 부(-)의 영향으로 유의하여, ESG 공시를 통한 책임 투자 활성화로 기업의 타인자본에의 의존도가 줄어들면서 공시가 재무 안정성에 도움이 될 것으로 보임
최신 ESG 정보공시 동향과 그린워싱	김동수(2022)	- ESG 워싱은 기업의 ESG 커뮤니케이션과 실제 실행 또는 성과 간의 차이에 의해 발생함 - ESG 정보공시의 핵심은 ESG 성과관리에서 시작해야 하며, 체계적 프로세스를 통해 정확한 정보가 공개될 수 있도록 ESG 커뮤니케이션에 대한 철저한 관리가 필요함
지속가능경영보고서의 공시와 기업가치	남해정(2023)	- 회계정보의 질이 우수할수록 자발적으로 비재무 정보를 공시함 - 자발적으로 지속가능경영보고서를 공시하는 기업들은 그렇지 않은 기업들에 비해 기업가치가 우수하게 나타남

12

선행연구

제목	저자(년도)	주요 내용
국민연금 기금 책임투자자와 외부 ESG 평가와의 관계에 관한 연구	정경화, 박삼복(2022)	- 대표적인 공적 연기금인 국민연금 기금이 5% 이상 지분을 보유하여 책임 투자를 고려한 항목으로 4년 연속 공시 대상에 포함된 종목이 1~3년 공시대상에 포함된 종목보다 ESG 평가 결과와의 관련성이 더 크게 유의하게 나타남 - 국민연금이 5% 이상 지분을 보유하여 책임투자를 고려한 투자종목으로 4년 연속 공시 대상에 포함된 종목 중 5년 연속 지분율을 확대한 종목이 지분율 증가 또는 감소로 보유 지분율의 증감이 있었던 종목보다 ESG 평가 결과와의 관련성이 더 크게 유의하게 나타남
최신 ESG 정보공시 동향과 그린워싱	김석수, 정재경(2023)	- 한국 제조기업의 ESG 평가등급은 총자산수익률에 양(+)의 요인을 끼치고 있으며 특히 지배구조 부분의 평가가 상대적으로 강하게 영향을 끼치고 있음

- 지속가능경영보고서 발간은 단기간의 수익성에는 부정적인 영향을 끼칠 수 있으나, 기업의 안정성 등 기업가치에 긍정적인 영향을 끼침
- ESG 워싱은 기업의 ESG 정보 공시와 실제 성과간의 차이에서 발생함
- ESG 평가 결과는 기업의 총 자산수익률이나 투자에 영향을 끼칠 수 있음
 - > 기존의 논문은 ESG 공시가 기업 가치에 미치는 영향과 ESG 평가 결과가 기업가치에 미치는 영향을 구분하여 분석
 - > 중대성 평가와 관련된 논문에서는 단일 업종에 대한 중대성 평가 방법 제시 및 결과에 대한 분석이 주를 이룸
 - > 본 논문에서는 ESG 공시와 ESG 성과의 경합성 ESG 등급에 미치는 영향을 알아보고
 - 업종별 중대 이슈의 특징과 중대성 평가의 개선방안에 대해 제시하고자 함

13

03

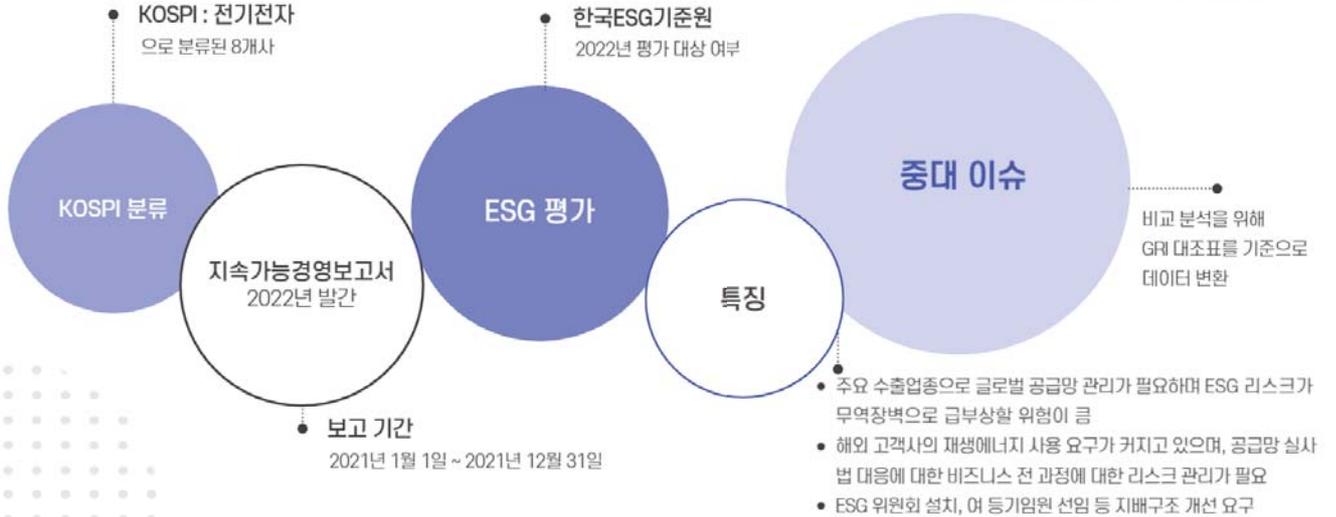
연구 방법

연구 절차도



연구대상

'배출', '에너지'가 중대 이슈 중 가장 많은 비중을 차지하였으며, '고용', '폐기물'이 그 뒤를 이음



연구대상

"고객안전보건", '배출', '에너지'가 중대 이슈 중 가장 많은 비중을 차지하였으며, '산업안전보건', '폐기물'이 그 뒤를 이음



연구데이터 분류 및 수집

구분	데이터 정보
수집 대상 기간	2021년 1월 1일 ~ 2021년 12월 31일
수집 정보	- 14개 기업별 지속가능경영보고서 - 14개 기업별 ESG 관련 뉴스기사 (검색어: 기업명 ESG) - 건설업종 ESG 관련 뉴스기사 - 14개 기업별 사업보고서 - 14개 기업별 트위터 (검색어: 기업명)
세부 수집 정보	- 기업별 중대성 평가 분석 결과 - 기업별 사업보고서 내 제재현황

TF-IDF

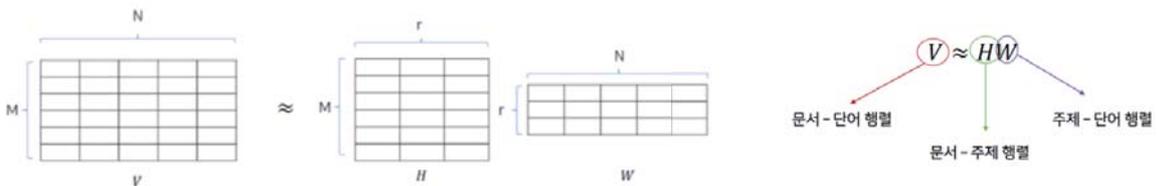
정보검색, 텍스트 마이닝 등에서 이용하는 가중치로, 문서 군에서 어떤 단어가 특정 문서 내에서 얼마나 중요한 것인지를 나타내는 통계적 수치

$$tfidf(t,d,D) = tf(t,d) \cdot idf(t,D)$$

어떠한 단어가 특정 문서 내에서 얼마나 중요한 것인지 = 데이터 문서군인 d에서 단어 t가 언급된 횟수 * 전체 데이터 수 / 단어 t가 포함된 데이터의 수

NMF(Non-negative matrix factorization)

'비음수 행렬 분해', '음수 미포함 행렬 분해'라고도 불리며, 명칭대로 음수를 포함하지 않는 행렬 H와 W의 곱으로 분해하는 알고리즘
행렬이 음수를 포함하지 않는 특성은 음수를 포함하지 않는 텍스트 분석을 용이하게 만들어 줌



$$F(A, B) = \|A - B\|^2 = \|V - HW\|^2$$

→ 최소값(0)으로 수렴할 때까지 감소 (A=B 가 되는 지점)

본 연구에서는 TF-IDF(문서-단어) 를 기반으로 (문서-주제(이슈))행렬, (주제(이슈)-단어) 행렬을 추출하게 되며, 해당 주제(이슈)의 개수는 기업의 지속가능경영보고서의 중대 이슈 개수와 동일하게 추출하여 비교가 용이하도록 진행

상관성 분석

두 변수 간에 어떤 선형적 관계를 가지고 있는지를 분석하는 방법으로, 두 변수는 서로 독립적인 관계이거나 상관된 관계일 수 있으며, 이 때 두 변수 간의 관계의 강도를 상관관계(Correlation, Correlation coefficient)라고 함

$$r_{XY} = \frac{\sum_i^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i^n (X_i - \bar{X})^2} \sqrt{\sum_i^n (Y_i - \bar{Y})^2}}$$

r 결과값	결과 해석
-1.0 ~ -0.7	강한 음적 선형관계
-0.7 ~ -0.3	뚜렷한 음적 선형관계
-0.3 ~ -0.1	약한 음적 선형관계
-0.1 ~ +0.1	거의 무시될 수 있는 선형관계
+0.1 ~ +0.3	약한 양적 선형관계
+0.3 ~ +0.7	뚜렷한 양적 선형관계
+0.7 ~ +1.0	강한 양적 선형관계

ESG공시와 ESG 성과의 정합성이 ESG 평가 등급에 미치는 영향을 분석하기 위해 한국 ESG 기준원의 등급표를 점수로 환산하여 정합성(%)과의 상관관계 분석을 진행함

04

분석 결과

NMF 분석 결과

NMF 분석결과로 도출된 중대 이슈 역시 GRI 대조표를 기준으로 데이터 변환 진행

A. 기업	NMF 결과
1	<p>유위험의 환경영 #기업 #기술 #환경영 #전지 #경제적 #가능성 #구조 #10대 ESG진행, ESG로 진정 "친환경 미래지향 진도"</p> <p>ESG위원회, 만들기는 쉽다... 제대로 하려면 어려운 문제다 [스캐일 리포트]</p> <p>10대 그룹, 'ESG 경영' 전담기구 상설기 설치</p> <p>전 세계 자산 규모 2년세 3배 ↑... 기업 투자 결정적 요소로 [상승기회]</p> <p>ESG진행, ESG 위원회 신설... "친환경 미래사</p>
2	<p>안전성 #사이보 #편중 #사업영 #중대 #경제적 #발생 #인공지능 #도덕적</p> <p>[스캐일 리포트] VR로 안전교육... 오봇이 현장점검...건설사 '안전'에 올인</p> <p>[중대세법 공보(1)]과불안적 '10대 건설사' CEO...누구?</p> <p>스마트폰·로봇계가 24시간 감시...기업들, 안전기술 아낌없이 투자</p> <p>현대중공업 "건설사고 줄일 위해 장구하다"...건설업체 '긴장'</p> <p>아웃부 감독도 소홀하다...로스코·GS건설, 근무와 안전 '빨간불'</p>
3	<p>여성성 #사이보 #편중 #사업영 #중대 #경제적 #발생 #인공지능 #도덕적</p> <p>건설업계 유망형</p> <p>@condine</p> <p>이로써</p> <p>@jyoung2</p> <p>ESG진행 첫 여성 사외이사에 초희원 변호사 선임</p> <p>건설업계 우수사은 유입... 여성이사회 선임·선사형 화두로</p> <p>대기업 CEO·여류수·ESG전문가...용해 사외이사 '대세' [스캐일 리포트]</p> <p>#노조 #편중 #사업영 #중대 #경제적 #발생 #인공지능 #도덕적 #사부</p> <p>[Fact]</p> <p>@gghung73</p> <p>초희원은 세상을 ...</p> <p>@hikim561</p> <p>Twitter #gptr</p>
4	



부문	키워드	NMF
E (4)	원재료(301)	#모탈 #주주채 #공명 #사업 #공사 #수주 #처리 #주인 #실계 #시장
	에너지(302)	#사업 #에너지 #배생 #투자 #수소 #친환경 #발전 #인공지능 #모탈 #시장
	배출(305)	#사업 #그린 #포탈 #당금 #물류 #물탄도 #센터 #친환경 #인중 #유입
S (3)	폐기물(306)	#스타트 #달렉스 #친환경 #생산 #사업 #그들 #발생 #태터리 #회장 #제일사
	사업장 안전보건(403)	#안전 #사고 #현장 #사업 #중대 #경제적 #발생 #근로자 #보건 #관리
	경사 및 단계교실의 자유(407)	#노조 #인노출 #발의 #방출건설 #기업 #인명 #사부적 #신하 #노조 #사부
G (3)	고급 안전보건(416)	#브랜드 #아파트 #공간 #주거 #입주인 #예보 #단지 #고급 #분양 #자이
	제품 및 서비스 라벨링 (417)	#위험의 #경영 #기업 #그들 #환경 #현시 #지속 #가능 #구조 #10대
	일반공시(102)	#여성 #사외이사 #선임 #편중 #사업영 #중대 #경제적 #발생 #인공지능 #도덕적 #사부
G (3)	일반공시(102)	#여성 #사외이사 #선임 #편중 #사업영 #중대 #경제적 #발생 #인공지능 #도덕적 #사부
	반부패(205)	#사건 #고발 #공명 #국방부 #군수 #기당행위 #수사 #거래 #자율 #대전

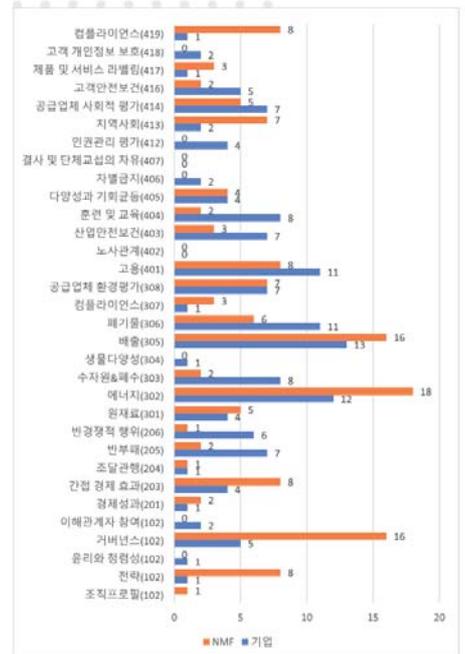
NMF 분석 결과 (전기 전자)

분석 결과, 거버넌스(102)에 대한 적합성 차이가 두드러지게 나타남
 현재 전기전자 내에서 여성 사외이사 선임 및 ESG 위원회 설치 등 지배구조 개선에 대한 요구가 반영된 것으로 보임

이어 전략(102), 컴플라이언스(102)가 NMF 결과에서 다빈도로 추출됨
 기업의 중대이슈 중 G(지배구조)의 비중이 현저히 낮고, 지배구조 부문의 중대이슈는 대체로 거버넌스(102) 혹은 반부패(205), 반경쟁적행위(206)에 집중되어 있음

인력 유지를 위한 고용(401)이 다빈도로 추출되었으며, 기업의 중대 이슈 대비, 지역사회에 대한 사회공헌이 활발하게 이루어지는 것을 확인할 수 있었음

수소 및 신재생에너지 사용과 기후변화 대응으로 에너지(302)와 배출(305)이 기업의 중대이슈, NMF 결과 모두 최다빈도로 추출됨

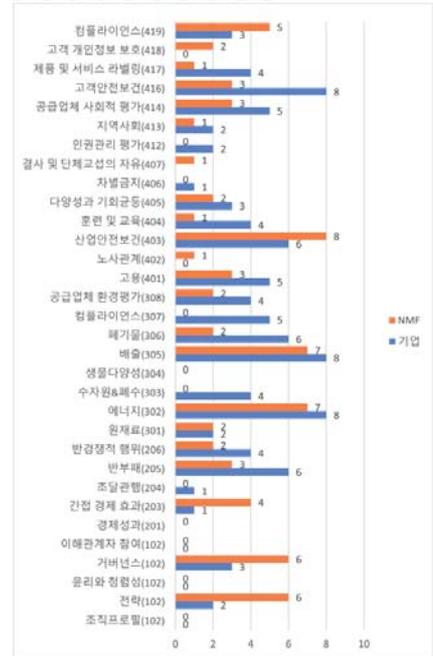


NMF 분석 결과 (건설업)

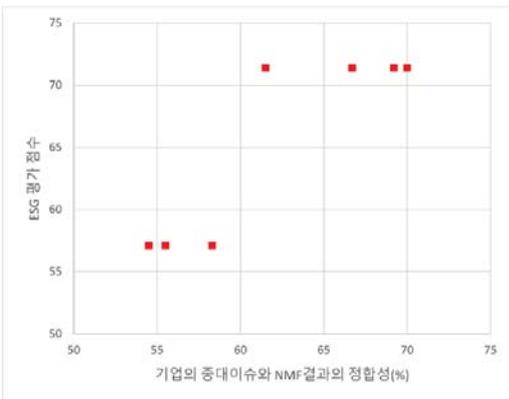
분석 결과, 컴플라이언스(307)에 대한 차이가 가장 크게 남
 건설업은 환경사고보다는 고용, 산업재해에 대한 사고 비율이 높기 때문에, 해당 결과에 대한 차이로 보임

이어 산업안전보건(403)이 다빈도 단어로 추출되었으며, 고객안전보건(416) 또한 큰 차이를 보임
 산업안전보건의 경우 산업안전보건법 위반이나 현장에서의 재해들로 인한 결과, 중대재해처벌법 도입에 대한 대비로 인한 성과들이 반영된 것으로 보임

환경부문에서는 수자원&폐수(303), 폐기물(306), 컴플라이언스(307)등이 실제 성과와는 다르게 다르게 나타남

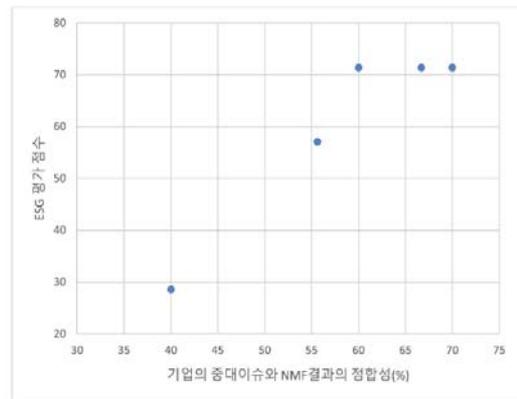


상관성 분석 결과 (전기전자)



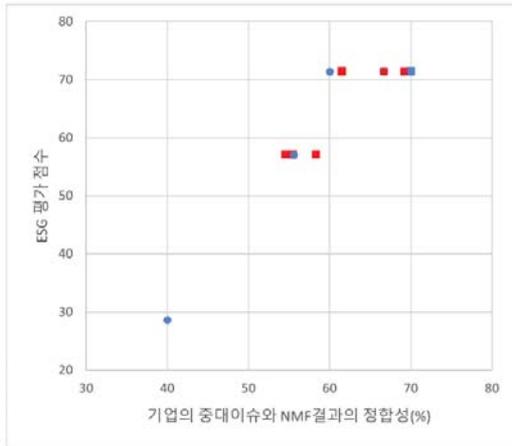
피어슨 상관계수: 0.8977666111843338
 p.value: 0.0024706409480213837
강한 양적 상관관계를 보이는 것으로 나타남

상관성 분석 결과 (건설업)



피어슨 상관계수: 0.9453754221871447
 p.value: 0.00439427112877951
강한 양적 상관관계를 보이는 것으로 나타남

상관성 분석 결과 (전기전자 + 건설업)



기업별 중대 이슈와 NMF 결과의 정합성과 ESG 등급 점수의 산점도를 살펴보면 선형에 가까운 산점도가 산출되었으며, 피어슨 상관계수가 0.9308048813825397, p.value가 1.3635167397249774e-06로 **강한 양적 상관관계**를 보이는 것으로 나타남

결론



- 전기전자는 환경부문에 적합한 대응을 하고 있으나, 중대 이슈에 지배구조 비중 확대 및 컴플라이언스 체제 확립 필요
- 건설업은 노동집약적인 특징을 지니기 때문에 산업재해와 관련된 민감한 이슈설정이 필요
- 반부패, 반경쟁적 행위, 컴플라이언스뿐만 아니라 ESG 경영체계에 대한 G (지배구조) 부문을 중대이슈에 포함한 기업의 ESG 등급이 높아지는 것을 확인

- 기업의 중대 이슈와 기업의 ESG 성과와의 정합성이 높을수록 기업의 ESG 등급이 높음을 확인

ESG 전략과 공시(커뮤니케이션) 간의 디커플링으로 인해 발생하는 '위상(위장행위)'이 ESG 평가등급 하락으로 이어질 수 있음

본 논문은 기업의 ESG 성과와 공시 간의 정합성이 ESG 평가로 이루어진다는 점을 확인하였으며, 이는 앞으로 지속가능경영보고서를 발간하는 기업들에게 정확한 ESG 공시를 통해 높은 ESG 등급을 받는 선순환 구조를 이끌어낼 수 있다는 점을 시사함 따라서 이슈의 성격과는 상관없이 성과와 공시가 일치하는 투명한 ESG경영이 이루어져야 함

현재 업종별 지속가능경영보고서 발간기업의 수가 적어 비교군이 부족하며, 글로벌 공급망에 대한 자료 확보가 어려움 각 기업은 자체 데이터를 활용하여 해당 분석을 진행한다면 국내외 평가기관에 대응할 수 있는 수단으로 활용가능할 것으로 보임

참고문헌

1. 관계부처 합동(2021), K-ESG 가이드라인 v1.0, 2021
2. 금융위원회, 금융감독원, 한국거래소(2021), 기업 부담은 줄이고, 투자자 보호는 강화하는 기업공시제도 종합 개선방안
3. 김동수(2022), 최신 ESG 정보공시 동향과 그린원성, 미래성장연구
4. 김석수, 정재경(2023), 한국ESG기준원의 ESG 평가가 재무성과에 미치는 영향에 관한 연구
5. 김이진, 류현정(2021), ESG평가체계 현황 분석 연구 환경영역(대)을 중심으로, 한국환경연구원
6. 관계부처 합동(2021), K-ESG 가이드라인 v1.0
7. 남해정(2023), 지속가능경영보고서의 공시와 기업가치, 한국보험학회
8. 박은수, 이지운(2022), 다학제적 관점의 ESG 중대성 평가를 통한 건설산업의 지속가능성 분석
9. 박지원, 이예림(2022), 국내외 ESG 평가사별 점수 비교: 국내 기업을 중심으로, 대외경제정책연구원
10. 박태양(2022), ESG경영 공시전환에 대응하는 중대도착 공시방법 연구-석유와 가스산업 중심으로, 한국산업경영시스템학회지, 45(1), 53-70
11. 변해영(2018), 비재무적 공시의 가치 관련성 연구: 지속가능경영보고서 공시를 중심으로, 윤리경영연구, 18(1), 175-204
12. 이나경, 임수영(2021), 산업분야에서의 ESG활용을 위한 기초적 연구 - 미국 시가총액 상위 5개 기업을 중심으로 -, KIEAE Journal, 21(3), 83-88
13. 이한울, 이지현, 이준희(2022), Transformer 기반의 토크 모델링을 이용한 지속가능경영보고서 분석, 한국정보처리학회 학술대회논문집, 29(1), 464-467
14. 이홍일(2020), 건설기업의 성공적 지속가능경영 추진 방안, 한국건설산업연구원
15. 이혜림, 류주현, 이승혜(2022), GRI 지표를 활용한 지속가능경영보고서의 비판적 분석, 한국회계학회 학술발표논문집, 2022(2), 1-15.
16. 정경화, 박삼복(2022), 국민연금 기금 책임투자외 외부 ESG 평가와의 관계에 관한 연구, 한국상업교육학회
17. 정명환(2009), 우리나라 기업의 환경정보 공시 현황과 제도적 개선방안, 대한경영학회지, 22(6), 3529-3561
18. 조문기, 김지영, 안형태(2018), 지속가능경영 보고를 위한 글로벌 표준- 지속가능경영보고서 발간 기업 비교 분석 -, 기술경영, 3(1), 1-26
19. 조찬희(2023), ESG 평가방법 비교: K-ESG 가이드라인을 중심으로
20. 차지미, 윤호근, 김미숙(2021), 지속가능경영보고서를 이용한 ESG 요소 정성적 분석, 한국정보과학회 학술발표논문집, 1753-1755
21. 최윤형, 이기호, 이상명(2022), 지속가능경영보고서의 중요성 분석을 통해 바라본 지속가능경영 이슈와 10년의 변화, Korea Business Review, 26(1), 125-148
22. 최신민, 윤천성(2022), 키워드 네트워크 분석을 활용한 ESG 연구 동향 분석, 인문사회 21, 13(2), 2073-2084
23. 한국거래소, ESG 정보 공개 가이드라인, 2021
24. Global Reporting Initiative(2021), GRI Standards GRI 3: Material Topics 2021
25. Jilke Gerstl, Karen Mass, Jeroen Sujs(2022), Materiality Assessment Is an Art, Not a Science: Selecting ESG Topics for Sustainability Reports, California Management Review 2022, Vol. 65(1) 64-90
26. KPMG(2020), KPMG Survey of Sustainability 2020

경청해주셔서
감사합니다

라이브 커머스 쇼핑이 소비자 구매의도에 미치는 영향

The Effect of Live Commerce Shopping on Consumer Purchase Intention

리우주(경북대학교 석사과정)
차오신(경북대학교 박사과정)
오세환(경북대학교 부교수)

1

목차

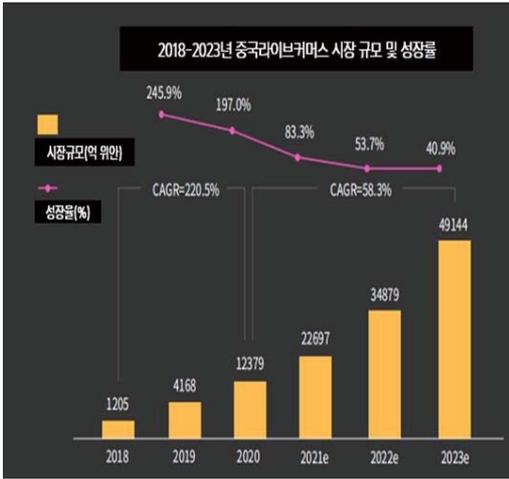
- 01 연구배경 및 동기
- 02 이론 배경
- 03 연구모형 및 가설
- 04 연구 방법론
- 05 시사점 및 한계점

2



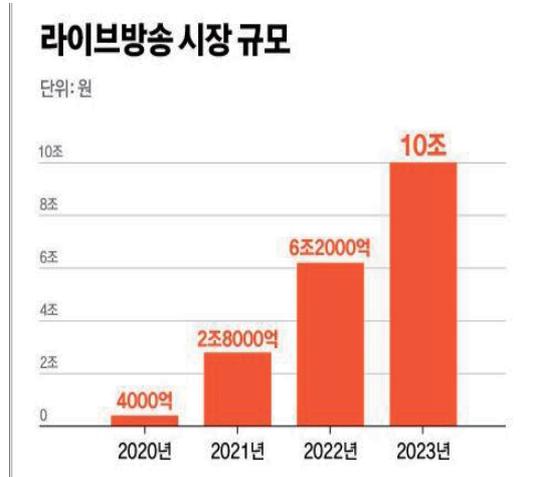
01 연구배경 및 동기

2018-2023년 중국 라이브 커머스 시장 규모



출처: <https://blog.naver.com/ggadduggi/222984776041>

2020-2023년 한국 라이브 커머스 시장 규모



출처: <https://blog.naver.com/webcaster/222440954316>

01 연구배경 및 동기



- 4차 산업혁명과 함께 네트워크의 대중화와 활용 범위는 점점 더 넓어지고 사람들의 일상 생활 곳곳으로 스며들고 있음.
- 전자 상거래도 시대의 요구에 따라 생겨나 날이 갈수록 강해지고 있으며, 그 중에 온라인 쇼핑은 전자 상거래의 중요한 구성 부분.
- 코로나-19 이후에 사람들의 소비 개념이 오프라인에서 온라인으로 이동함에 따라 유망한 시장이라는 장점이 있지만 치열한 시장 경쟁과 복잡하고 다양한 프로젝트와 같은 도전에 직면하고 있음.
- 동시에 라이브 쇼핑의 발전은 경제 회복을 촉진할 새로운 기회를 제공.





02

이론 배경

2.1 The S-O-R framework



- Skinner(1935): 환경과 자극의 관계를 연구하여 환경이 인간의 행동에 직접적으로 영향을 미칠 수 있음을 증명했음.
- Kotler(1999): 소비자 행동 모델을 처음으로 제안하여 마케팅 분야에 이론을 적용하고 소비자 구매 행동에 대한 심층 연구를 수행했음.
- SOR 모델: 사람들의 내적 감정과 인지적 요인에 초점을 맞추고, 소비자 행동 의도를 연구하는 데 적합함.
- Donovan & Rossiter(1994): 소매 환경에 처음으로 SOR 모델을 적용했으며 주로 소매에서 다양한 환경 자극이 소비자 구매 행동에 미치는 영향을 연구했음.

5

02

이론 배경

2.2 Live Commerce

- 라이브 커머스: 다양한 소셜 커머스 특성 뿐만 아니라 독특한 미디어 특성 포함 (소셜 커머스: 소셜 미디어를 매개로 한 상업적인 활동을 의미).
- Li Yuhong(2017): 소셜 이벤트가 발생하고 발전함에 따라 기업은 라이브 방송 플랫폼을 사용하여 브랜드를 노출하거나 제품 판매를 늘린 것.
- Cai et al. (2018): 라이브 커머스를 소셜 커머스의 한 형태로 분류하며, 새로운 형태의 쇼핑 방식으로 소개.

6





02

이론 배경

2.3 Features of Live commerce

➤ Price concessions → Utilitarian Value, Hedonic Value

- Smith & Colgate(2007): 소비자에게 적합한 경험, 느낌, 감성을 만들어내는 제품은 실용적인 가치와 쾌락적 가치와 밀접한 관계가 있다고 실증하였다.
- Cachero-Martínez & Vázquez-Casielles(2017): 가격, 제품 및 할인 같은 실용적인 경험은 실용적 가치를 증가시키는 중요한 요소이다.
- 변광인 외(2013) 연구에서 가격 할인 혜택에 대하여 소비자가 만족할수록 가치가 높아지고 가격 할인혜택이 쇼핑을 즐겁게 하기때문이라고 하였다.
- H1a: 가격 할인은 실용적 가치에 정(+)의 영향을 미칠 것이다.
- H1b: 가격 할인은 쾌락적 가치에 정(+)의 영향을 미칠 것이다.

7

02

이론 배경

2.3 Features of Live commerce

➤ Real time review → Utilitarian Value, Hedonic Value

- Zhang et al.(2022): 실시간 온라인 리뷰에 참여함으로써 소비자는 강력한 사회적 존재감에 몰입하고 소비자의 실용적 가치와 구매 결정에 강력한 영향을 미친다.
- Ham, Juyeon, et al. (2019): 논리적이고 진실된 리뷰는 온라인 리뷰의 쾌락적 가치를 향상시킬 수 있다.
- H2a: 실시간 리뷰는 실용적 가치에 정(+)의 영향을 미칠 것이다.
- H2b: 실시간 리뷰는 쾌락적 가치에 정(+)의 영향을 미칠 것이다.

8





02

이론 배경

2.4 Features of Internet Celebrities

➤ Interactivity

- 상호작용성은 사람들이 정보를 교환할 수 있는 정도를 말한다(Tajvidi, M.; Wang, Y.,2021; Huang, M.-H,2003).
 - 제품에 대한 즉각적인 피드백과 응답은 시청자들의 제품과의 상호작용을 증가시키고, 제품을 평가할 능력에 대한 인식을 높인다(J. Steuer,1992).
 - 이명수(2000): 즐거움은 상호작용을 통하여 유발될 수 있으므로 상호작용성이 높을수록 오락성이 높다고 지각할 가능성이 높다.
- H3a: 상호작용성은 가격 할인성과 실용적 가치 사이의 관계를 조절할 것이다.
 - H3b: 상호작용성은 가격 할인성과 쾌락적 가치 사이의 관계를 조절할 것이다.
 - H4a: 상호작용성은 실시간 리뷰와 실용적 가치 사이의 관계를 조절할 것이다.
 - H4b: 상호작용성은 실시간 리뷰와 쾌락적 가치 사이의 관계를 조절할 것이다.

9

02

이론 배경

2.4 Features of Internet Celebrities



➤ Social Presence

- 라이브 커머스 사회적 존재감은 구매자와 판매자 사이의 심리적 거리를 줄여 제품과 인터넷 유명인에 대한 소비자의 신뢰를 높일 것이다(Darke et al., 2016).
 - 라이브 커머스는 자세한 시연을 위해 제품을 직접 사용해 보는 경우가 많으며, 일반적으로 제품 경험에 대한 생생한 표현이 수반된다. 이는 소비자로서 하여금 제품을 사용하는 것이 어떤 것인지 상상하게 하고 이후에 정서적 반응을 생성하도록 유도할 수 있다 (Y. Chen et al., 2018).
- H3c: 사회적 존재는 가격 할인성과 실용적 가치 사이의 관계를 조절할 것이다.
 - H3d: 사회적 존재는 가격 할인성과 쾌락적 가치 사이의 관계를 조절할 것이다.
 - H4c: 사회적 존재는 실시간 리뷰와 실용적 가치 사이의 관계를 조절할 것이다.
 - H4d: 사회적 존재는 실시간 리뷰와 쾌락적 가치 사이의 관계를 조절할 것이다.

10



02

이론 배경

2.5 Utilitarian Value

- 실용적 가치: 소비자가 효율적인 쇼핑 과정으로 인지하는 쇼핑 동기 중 하나, 소비자가 온라인으로 구매할 때 중요하고 합리적인 결정을 내리는 것으로 정의(Babin, Darden, & Griffin, 1994; Hirschman & Holbrook, 1982).
- 온라인 쇼핑 플랫폼이나 기타 소셜 사이트에서 제공하는 기능적 이점으로서, 사용 편의성과 가격 정보 등을 포함 (Hsu & Lin, 2016).
- 사용의 용이성, 제품 정보, 제품 다양성과 같은 요소들이 온라인 구매 의도에 영향을 미치는 것을 보여줌(Sramova & Pavelka, 2019; Wolfenbarger & Gilly, 2001).

11

02

이론 배경

2.6 Hedonic Value

- 쾌락적 가치: 쇼핑의 즐거움과 같은 경험적인 쇼핑 행동을 특징으로 하는 소비자의 선호도라고 정의(Delafruz et al., 2009); 좋은 느낌과 즐거움과 같은 이익적인 측면이 아닌 것들의 기능으로 정의(Heijden, 2004).
- 인지된 즐거움은 일부 플랫폼에서 온라인 구매 시 즐거움을 결정하기 위해 첫 번째 변수로 선택됐고, 플랫폼 사용자가 느끼는 즐거움 수준이 높을수록 그들은 더 많은 쇼핑 시간을 갖고 온라인 서핑을 할 것이다(Mayada & Ahmed, 2019).

12





02 이론 배경

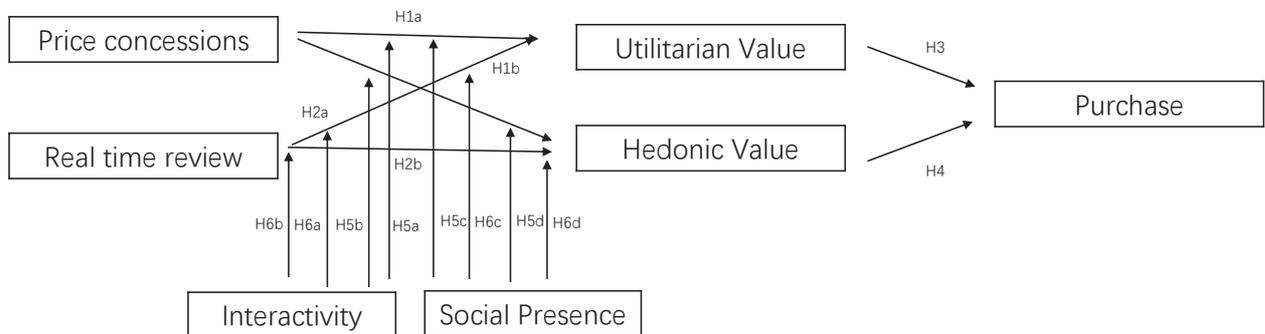
2.7 Purchase

- Babin et al.(1994): 실용적 쇼핑 가치와 쾌락적 쇼핑 가치의 존재를 검증하고, 소비금액과 만족도 등 소비자의 구매의도에 영향을 줄 수 있다.
 - Wu et al.(2019): 소비자는 라이브 쇼핑 목표 달성을 희망할 뿐만 아니라 쇼핑 과정에서 좋은 경험을 얻고 레저와 오락을 즐기며 소비자가 진심으로 만족한다고 생각하기를 희망한다.
 - Ashfaq(2019): 소비자가 쇼핑 과정에서 높은 수준의 즐거움을 느낄 때 소비자 만족도가 높아지고 소비자의 구매 의도가 높아짐.
- H5: 실용적 가치는 구매의도에 정(+)의 영향을 미칠 것이다.
 - H6: 쾌락적 가치는 구매의도에 정(+)의 영향을 미칠 것이다.

03 연구모형 및 가설



Features of Live commerce



Features of Internet Celebrities



04 연구 방법

- 조사 방법: 설문조사
- 데이터 분석 방법: Smart PLS 4
- 2023.06, 국내외 온라인 라이브 쇼핑 소비자 44명
- Likert 5점 척도 사용

<표 1> 신뢰성 및 타당성 분석 결과

요인	요인 적재 값	Cronbach's α	CR	AVE	VIF	
Price Concessions	PC1	0.916	0.885	0.928	0.811	3.031
	PC2	0.881				2.574
	PC3	0.905				2.258
Real-time Review	RTR1	0.786	0.804	0.886	0.722	1.662
	RTR2	0.921				2.837
	RTR3	0.836				2.041
Interactive	I1	0.856	0.865	0.917	0.788	1.991
	I2	0.908				2.708
	I3	0.898				2.328
Social Presence	SP1	0.890	0.878	0.925	0.805	2.447
	SP2	0.922				2.994
	SP3	0.878				2.180
Utilitarian Value	UV1	0.908	0.850	0.909	0.769	2.506
	UV2	0.889				2.281
	UV3	0.832				1.784
Hedonic Value	HV1	0.940	0.745	0.881	0.788	1.544
	HV2	0.832				1.544
Purchase	P1	0.898	0.825	0.898	0.748	4.596
	P2	0.965				5.848
	P3	0.713				1.634

15

05 시사점 및 한계점

시사점

1. 라이브 커머스 분야에서 소비자의 구매의도에 관한 연구를 확장하는 점이다. 연구 결과는 라이브 커머스 분야에 영향과 도움을 줄 수 있는 이론적 근거를 제공할 수 있을 것으로 기대한다.
2. 라이브 커머스는 소비자의 지각된 가치를 향상시킬 수 있다. 소비자의 구매 의도를 자극할 뿐만 아니라 라이브 커머스 과정에서 전통 전자상거래와의 다른 쇼핑의 재미를 경험할 수 있다.

16





05 시사점 및 한계점

시사점

3. 라이브 커머스에서의 인터넷 유명인(Internet Celebrities)특성을 확인할 수 있다. 일반적인 구매 과정보다 라이브 커머스 중 소비자의 라이브 댓글이나 의 중요한 상품 설명은 종종 소비자의 구매 의도에 영향을 미친 것이다. 라이브 쇼핑 분야에서 소비자와의 상호작용, 인터넷 유명인(Internet Celebrities)과의 상호작용은 소비자의 구매 의도를 결정짓는 요인이다.

한계점

- 데이터 한계
- 플랫폼의 한계
- 지역의 한계

17



감사합니다

18

The Price Formation of the Emerging Digital Art Market: Entry Timing and Market Experience of NFT Users

발표자: 김우경 / 숙명여대
한국경영정보학회 융합학술대회
2023.08.16

Agenda

- 1) Motivation & Purpose
- 2) Theory & Hypothesis
- 3) Data
- 4) Results
- 5) Conclusion & Future Work

Motivation

- NFT art market has two institutional logics - Investment logic / Art logic (Lounsbury, 2007).

- 1) **Investment logic** – NFT as an asset
- 2) **Art logic** – NFT as an art

- The NFT itself is also very high in that it is unclear whether it will be defined as an asset with financial value or an artwork subject to appreciation.
- The NFT market is lack of indicators, information, and categorized systems for market participants to use for decision-making.
- Few studies have fully considered the uniqueness of this market.

-> **It is difficult to find consistent standards of price and market participants need them to behave in the market.**

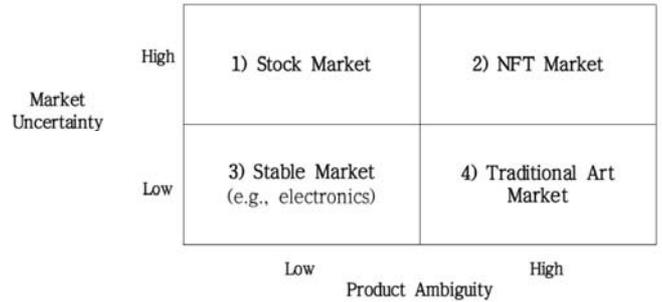


Figure 1. Four types of markets by market uncertainty and product ambiguity

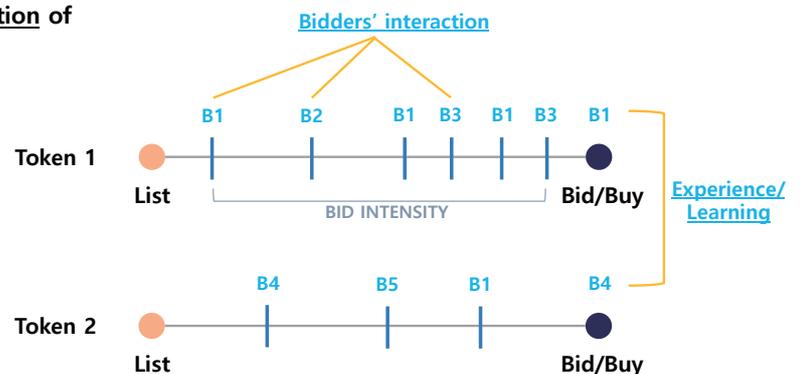
Purpose

**What are the factors that influence the listing(by creator)/selling(by bidder) price of each token?*

- This study assumes that **NFT price formation is a combination of**

- 1) The impact of external factors
- 2) Differential patterns of market players' behavior
- 3) Experience / learning
- 4) Interaction intensity

is analyzed.



Theory & Hypothesis – (1)

- When deciding whether to buy or sell a stock, for instance, people often consult the price movements of stocks that belong to the same stock category (Zuckerman, 1999).
- Although these volatile patterns of market movement are widely investigated in various fields of research such as **herding behavior theory, fad and fashion research, and collective action theory** (Banerjee, 1992; David & Strang, 2006; Strang & Macy, 2001; Granovetter, 1978), the recent development in complexity theory also sheds additional light on why such rapid “**phase transition**” occurs (Chiles, Meyer, & Hench, 2004).



H1-1: As the Ethereum price increases, the list price decreases.

H1-2: As the Ethereum price increases, the selling price increases.

Theory & Hypothesis – (2)

- A central tenet of bounded rationality perspective is that “decision makers do not attend to all of their goals at once, nor all of their alternatives, nor all of the consequences of the alternatives” (March & Simon, 1993, p.4).
- Numerous studies found that **an organization under uncertain market environments is likely to attend to the moves of other organizations**, especially those that are bigger than or similar to the focal organization (Haveman, 1993; Haunschild & Miner, 1997; Schwab, 2007).



H2-1: Early entrants in the market tend to set higher list prices.

H2-2: Early entrants in the market tend to have higher selling prices.

H2-3: Early entrants in the market set higher list price when Ethereum price increases.

H2-4: Later entrants in the market have lower selling price when Ethereum price increases.

Theory & Hypothesis – (3)

- If such cues are unavailable, they can engage in **learning from past experience of their own and others**, exhibiting path-dependent behaviors (Gavetti, Levinthal, & Ocasio, 2007; Lampel, Shamsie, & Shapira, 2009).
- **One such information may come from actions of others**: given the ambiguous categorization of the product, actors attend to how other actors behave, thereby making inference about the quality and expected value of the product (Haunschild & Miner, 1997; Baum et al., 2000).
- In the NFT market, **it can be expected market players to learn from their experience or others' experience(performance) and update standards from them**(White, 2009).

 **H3-1: Users with a higher success rate in the past tend to set higher list prices.**

H3-2: Users with a higher success rate in the past tend to have higher selling prices.

H3-3: As the average market selling price increases, the list price tends to be set higher.

H3-4: As the average market selling price increases, the selling price tends to be higher.

Theory & Hypothesis – (4)

- Instead, they restrict their attention to **easily observable and highly salient information**, and use this selected information as a signal for an appropriate response action (Spence, 1974; Nelson, 1974; Feldman & March, 1981; Hoffman & Ocasio, 2001).
- Individuals are embedded in a web of relationships that are formed via **repeated interaction over time** in an online community (Samar, 2011).

 **H4-1: As there is more interaction among bidders, the selling price tends to be higher.**

Data

- **We will consider following groups of variables**
 - ① Market Conditions
 - ② Creator Characteristics
 - ③ Bidder Characteristics
 - ④ Creator-Bidder Characteristics
- **Dataset**
 - Token-wise data for the analysis of selling prices (1 row per token / daily & weekly data)
 - 363,920 mint/list/bid history, 38 variables
 - Data from *Foundation* (<https://foundation.app/>)

Data

- **Key Variables**
 - **Control variables:**
 - listing price,
 - days elapsed, days elapsed*N. of bids
 - N. of bids, concentration(t-1), N. of tokens minted, N.of tokens sold(t-1),
 - prior N. of tokens(creator), Prior N. of tokens(bidder),
 - ethereum price growth rate,
 - early adopter, early adopter*Ethereum price gr. rt.,
 - prior success rate(creator), Listing_price*prior success rate(bidder)
 - **Dependent variables:** logged
 - listing price
 - selling price
 - **Independent variables:**
 - The half of the initial market participants is classified as early, and the remaining 50% is classified as late entrants.

Data

Figure 2. Monthly total sales

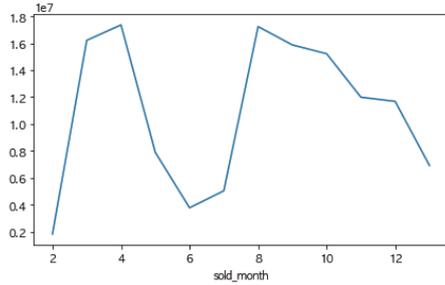


Figure 3. Monthly creators sold tokens

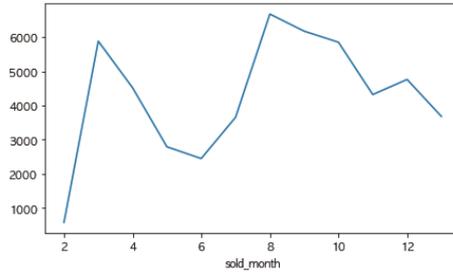


Figure 4. Monthly average of listing price (USD)

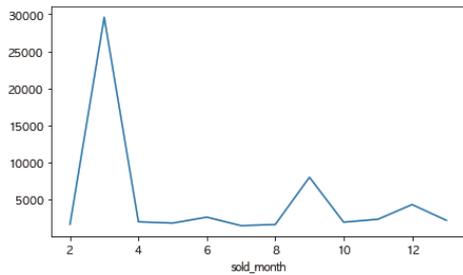
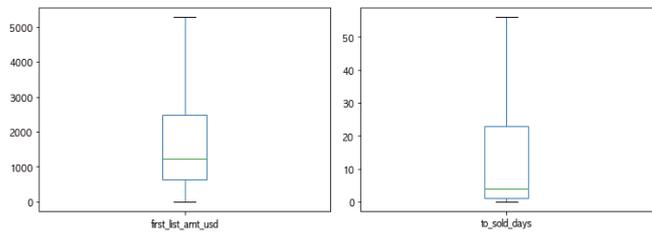


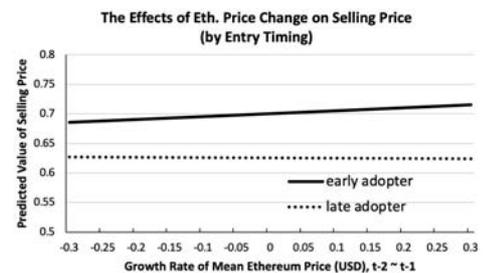
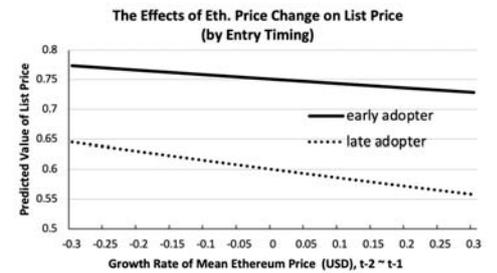
Figure 5-6. Listing price (USD) / Days to sold



Results (1)

Table 1. Mixed Effect Models of the Token Listing Price

Variables	Model 1		Model 2		Model 3		Model 4	
	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
constant	-0.505	0.010 ***	-0.511	0.010 ***	-0.557	0.011 ***	-0.598	0.011 ***
concentration (t-1) /1000	0.011	0.002 ***	0.015	0.002 ***	0.003	0.002	0.004	0.002
N. of tokens minted (t-1) /1000	0.008	0.003 *	-0.006	0.003	-0.023	0.003 ***	-0.041	0.004 ***
prior N. of tokens (creator) /100	0.121	0.021 ***	0.117	0.022 ***	0.086	0.022 ***	0.183	0.022 ***
ethereum price growth rate (t-1)			-0.134	0.018 ***	-0.244	0.031 ***	-0.302	0.032 ***
early adopter					0.224	0.009 ***	0.231	0.009 ***
early adopter × ethereum price gr. rt.					0.144	0.039 ***	0.146	0.038 ***
prior success rate (creator)							0.048	0.002 ***
N. of tokens sold (t-1) / 1000							0.072	0.008 ***
variance (creator)	0.375	0.004	0.374	0.004	0.364	0.004	0.368	0.004
variance (e)	0.389	0.002	0.389	0.002	0.388	0.002	0.384	0.002
Log-Likelihood	-137482.220		-133735.390		-137002.390		-136500.850	
Chi-Square vs. Linear Model	44371.020		44290.180		43793.230		44591.080	



Results (1)

H1-1: As the Ethereum price increases, the list price decreases.

➔ Supported

H2-1: Early entrants in the market tend to set higher list prices.

➔ Supported

H2-3: Early entrants in the market set higher list price when Ethereum price increases.

H3-1: Users with a higher success rate in the past tend to set higher list prices.

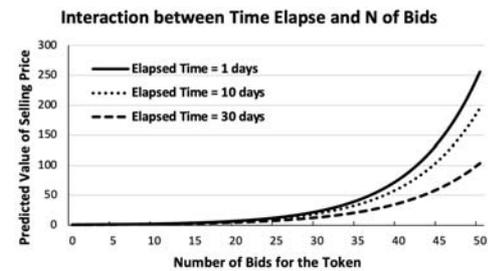
➔ Supported

H3-3: As the average market selling price increases, the list price tends to be set higher.

Results (2)

Table 2. Mixed Effect Model of the Token Selling Price

Variables	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coeff.	SE								
constant	-0.438	0.010 ***	-0.436	0.010 ***	-0.469	0.010 ***	-0.516	0.011 ***	-0.525	0.011 ***
listing price (log)	0.635	0.003 ***	0.635	0.003 ***	0.628	0.003 ***	0.637	0.005 ***	0.640	0.005 ***
days elapsed / 100	-0.240	0.005 ***	-0.240	0.005 ***	-0.244	0.005 ***	-0.244	0.005 ***	-0.162	0.006 ***
N. of bids (token)	0.117	0.001 ***	0.117	0.001 ***	0.116	0.001 ***	0.117	0.001 ***	0.122	0.001 ***
concentration (t-1) / 1000	-0.001	0.002	-0.002	0.002	-0.010	0.002 ***	-0.007	0.002 **	-0.007	0.002 **
N. of tokens minted (t-1) / 1000	-0.005	0.003 +	-0.006	0.003 *	-0.015	0.003 ***	-0.032	0.004 ***	-0.030	0.004 ***
prior N. of tokens (creator) / 100	0.177	0.019 ***	0.176	0.019 ***	0.156	0.019 ***	0.109	0.019 ***	0.112	0.019 ***
prior N. of tokens (bidder) / 100	0.012	0.001 ***	0.012	0.001 ***	0.013	0.001 ***	0.011	0.001 ***	0.011	0.001 ***
ethereum price growth rate (t-1)			0.052	0.018 *	-0.007	0.031	-0.079	0.032 *	-0.077	0.032 *
early adopter					0.113	0.007 ***	0.125	0.007 ***	0.123	0.007 ***
early adopter*ethereum price gr. rt.					0.078	0.038 *	0.099	0.038 **	0.102	0.038 **
prior success rate (creator)							-0.010	0.004 **	-0.009	0.004 *
listing_price*prior success rate (creator)							-0.045	0.004 ***	-0.044	0.004 ***
prior success rate (bidder)							0.079	0.009 ***	0.076	0.009 ***
listing_price*prior success rate (bidder)							0.020	0.008 *	0.019	0.008 *
N. of tokens sold (t-1) / 1000							0.065	0.007 ***	0.062	0.007 ***
days elapsed*N. of bids / 100									-0.057	0.003 ***
variance (creator)	0.090	0.002	0.090	0.002	0.089	0.002	0.089	0.002	0.087	0.002
variance (e)	0.154	0.001	0.154	0.001	0.153	0.001	0.152	0.001	0.151	0.001
Log-Likelihood	-31323.534		-31258.653		-31126.101		-30943.954		-30759.461	
Chi-Square vs. Linear Model	9092.340		9054.630		9234.780		9415.430		9225.830	



Results (2)

H1-2: As the Ethereum price increases, the selling price increases.	➔	Supported
H2-2: Early entrants in the market tend to have higher selling prices. H2-4: Later entrants in the market have lower selling price when Ethereum price increases.	➔	Supported
H3-2: Users with a higher success rate in the past tend to have higher selling prices. H3-4: As the average market selling price increases, the selling price tends to be higher.	➔	Supported
H4-1: As there is more interaction among bidders, the selling price tends to be higher.	➔	Supported

Conclusions & Future Work

- NFT market has new and unique characteristics, allowing market participants to hardly find consistent standards of price and it is difficult to make decisions on in the market.
- Market players need reliable information to behave for their purpose to engage in the market.
- We found that NFT price formation is influenced by external factors, differential patterns of market player's behavior, their own or others experience, and interaction.
- Furthermore, it can be considered how individual actors embed different logic and participate in the market, and how they decide their roles as buyers and creators accordingly, and how this role experience affects performing the following roles for future work.

건강 기능성 게임 수용에 영향을 주는 요인에 관한 연구

A Study on the Factors Influencing on the Adoption of Serious Games for Healthcare

2023. 8.

건국대학교 글로벌캠퍼스 경영학과

김용영



1. 연구 배경 및 목적

연구 배경 및 목적

■ 디지털 헬스케어에 대한 관심 증대

- COVID-19 감염의 세계적 유행으로 의료접근성이 제한되고, 인구 고령화 및 생활 수준의 향상으로 인해 만성질환 환자의 의료 수요가 증가
- 인공지능, 빅데이터 등 정보통신기술이 발전함에 따라, 약물이나 수술 대신 모바일 어플리케이션이나 웨어러블 기기, 기능성 게임, 가상현실 등 디지털 기기를 통해 질병을 치료

■ 디지털 치료제 Digital Therapeutics, 이하 'DTx'

- 환자의 치료 목적으로 모바일 어플리케이션, 가상/증강 현실, 기능성 게임, 인공지능 등을 이용하여 치료 효과를 내는 소프트웨어 기술
- 전 세계 DTx 시장은 2021년 31억 달러에서 매년 연평균 성장률이 31.4% 증가하여 2025년에는 131억 달러에 이를 것으로 예상
- 2022년 말 현재 미국 Food and Drug Administration (FDA)에서 인정받은 디지털 치료제는 60개 이상

3

연구 배경 및 목적

■ 페어 테라퓨틱스 Pear Therapeutics, 이하 '페어'

- 아래 3종을 포함하여 DTx를 선도하는 기업
- 2017년 미국 FDA에서 최초로 승인 받은 약물중독 치료 DTx → 리셋 reSET
- 아편유사제 사용 장애 환자를 인지행동요법(CBT)을 통해 치료 → 리셋-O reSET-O
- 우울을 동반한 만성 불면증을 치료하기 위한 DTx → 솜리스트 Somryst
- 최근 자금 압박 중인 페어의 매각 추진(2023.03) → 4월 나스닥 거래소로부터 상장폐지 통보

■ 페어가 어려운 상황에 처한 다양한 원인

- DTx를 의료진과 환자에게 익숙하게 하는데 부족, 새로운 치료 방법인 DTx를 시장에 침투할 수 있도록 보험사 등 다양한 이해관계자 설득에 고전
- **치료제의 한계**: DTx는 의사의 처방을 받아야 하는 치료제이기 때문에 자발적 수용 제약
 - FDA의 승인은 현재 치료제의 유통 경로를 따라야 한다는 점을 전제
 - 주의력 또는 결합형 ADHD가 있는 8-12세 어린이의 주의력을 향상시키는 FDA 승인 비디오 게임 치료제인 아킬리 Akili사의 엔데버알엑스 EndeavourRx도 의료진의 처방이 없이는 이용 불가

4

연구 배경 및 목적

■ 건강 기능성 게임

- 치료제가 아닌 디지털 웰니스 제품의 성격을 갖춘, 즉 FDA나 식품의약품안전처(이하 '식약처')의 승인을 받지 않고 건강증진을 목적으로 개발된 게임
- DTx와 동일하게 건강 기능성 게임도 24시간 모니터링이 가능하고, 수집된 데이터를 기반으로 맞춤형 게임을 제공할 수 있다는 특징이 동일
- DTx와 같이 다양한 혜택을 주는 정보통신기술 기반의 신제품이나 신서비스가 개발되고 출시되더라도 사용자(또는 환자)가 수용하여 사용하지 않는다면 이는 실패로 이어짐
- 건강 기능성 게임의 경우 건강의 증진이라는 혜택도 있지만, 한편 건강을 해칠 수 있는 부작용 발생 가능 → 일반적으로 건강 관련 기술은 혜택과 위험이 공존

5

연구 배경 및 목적

■ 연구 목적

- 인지된 혜택과 위험을 동시 고려 가능한 밸런스 프레임워크 valence framework를 기반으로 사용자의 건강 기능성 게임에 대한 수용 의사결정에 영향을 주는 촉진/저해 요인을 파악
- 선행연구를 검토하여 건강 기능성 게임 수용자의 인지된 혜택과 인지된 위험을 측정할 수 있는 척도를 개발하고 타당화

■ 본 연구의 구체적 목적

- 밸런스 프레임워크를 바탕으로 건강 기능성 게임 수용에 있어서 인지된 혜택과 인지된 위험을 동시에 고려하는 연구모형을 개발
- 문헌연구를 바탕으로 인지된 혜택과 인지된 위험의 설문 항목을 도출하여 검증
- 건강 기능성 게임의 수용을 촉진하는 요인과 저해하는 요인을 밝혀 웰니스 제품인 건강 기능성 게임이 확산될 수 있는 정책 방안이나 마케팅 방안을 모색할 수 있는 기초 자료 제공

6



2. 문헌 연구

건강 기능성 게임

■ 기능성 게임 serious game

- 1977년 사회학자 Clark C. Abt의 저서 <기능성 게임(Serious Game)>에서 보드게임으로 의사결정 교육의 성과를 높이기 위한 시도로 시작
- Abt는 ‘**Serious Game**’을 ‘사용자에게 놀이와 즐거움이 주된 목적이 아닌 교육에 주목적을 두고 있는 ‘진지한 serious’ 게임’으로 정의
- 게임의 재미 요소에 특별한 목적을 부가하여 제작된 게임으로 ‘Serious Game’과 ‘Functional Game’을 혼용해서 사용
- 초창기 기능성 게임은 주로 군사용으로 만들어졌으나 현재에는 의료건강 분야, 교육, 전문가 훈련, 정치사회적 인식 개선 등 적용 범위가 다방면으로 확장

■ 건강 기능성 게임

- 의료 및 건강관리 분야에 기능성 게임을 응용하여 적용하는 현상은 지속해서 성장
- 최근에서 의료를 건강과 치료를 포함하는 개념으로 보고 “건강”이라는 것은 추상적인 관점에서 모든 다양한 영역의 자원으로써 모든 종류의 게임에 접근 가능

디지털 헬스 산업의 범주화 [Dang et al., 2020]

- **디지털 헬스** digital health
 - 정보기술과 헬스케어를 융합한 넓은 개념으로 미국 FDA에서는 디지털 헬스의 범주를 모바일 헬스, 건강정보기술, 웨어러블 기기, 원격 의료와 원격진료, 개인맞춤형 의료를 포함
- **디지털 의학** digital medicine
 - 건강을 측정하는 근거 기반의 소프트웨어 또는 하드웨어로 ‘의학적 근거’가 필요하며, ‘다양한 규제’가 뒤따름
 - 디지털 의학 제품이 의료 기기로 분류되는 경우 승인이 필요. 다른 의약품, 기기 또는 의료 제품을 개발하기 위한 도구로 사용되는 디지털 의약품은 해당 심사 부서의 규제 승인 필요
- **디지털 치료제** digital therapeutics, DTx
 - 환자의 치료 목적으로 모바일 어플리케이션, 가상·증강 현실, 게임, 인공지능 등을 이용하여 치료 효과를 내는 소프트웨어 기술
 - ‘의료기기’로 분류되기 때문에 검증을 위한 ‘임상 근거 및 관련 규제’ 필요

9

디지털 헬스 산업의 범주화: SiMD vs. SaMD

- **의료용 소프트웨어 의료기기 분류: SiMD vs. SaMD**
 - Software in a Medical Device (SiMD)
 - 하드웨어 형태의 기존 의료기기에 탑재되어 단독으로 의료 목적을 수행할 수 없는 소프트웨어 프로그램으로, 의료기기를 운영하고 이를 통해 얻어진 정보를 처리하는 소프트웨어 기기
 - Software as a Medical Device (SaMD)
 - 하드웨어 의료기기의 일부가 아닌 하나 이상의 의료 목적으로 사용하기 위한 소프트웨어 의료기기이며, DTx는 SaMD로 분류

10

디지털 헬스 산업의 범주화: 디지털 웰니스

■ 디지털 웰니스 digital wellness

- 디지털 웰니스의 일부로 분류 가능. 그러나 Dang et al.(2020)이 제시한 디지털 헬스에는 체중 감량 지원 앱, 요가 학습 앱, 달리기 지도 앱 같은 일반적인 건강을 위해 소비자가 다운로드할 수 있는 앱 스토어의 iOS 및 안드로이드 앱과 같은 제품은 미포함
- Kappe(2022)는 이러한 디지털 헬스 산업 범주화의 문제점을 해소하고자 디지털 웰니스와 디지털 의약 사이에 디지털 웰니스 범주를 추가
- 디지털 웰니스 앱을 이용하여 사용자의 체중을 관리하거나, 영양 섭취를 최적화하거나, 휴대전화를 사용하는 시간을 줄여 사용자의 건강 향상 가능
- 일반적으로 디지털 웰니스 제품은 안전성이나 효과 입증 불필요
- 건강 기능성 게임도 디지털 웰니스 제품으로 분류되어 사용자가 자율적으로 선택하여 사용할 수 있으며, DTx와 같은 규제를 받지 않기 때문에 유통도 자유롭다는 특징
- DTx로 활용이 필요할 경우 FDA나 식약처에서 요구하는 안정성과 효과를 입증하여 DTx로 승인받고, 의료진이 환자에서 처방할 수 있는 장점

밸런스 프레임워크

■ 인지밸런스이론 cognitive valence theory

- 개인과 개인 사이의 인과관계 형성에 있어 긍정적 심리 요인과 부정적 심리 요인이 개인의 인간관계 형성에 영향을 미친다는 점 설명 [Andersen, 1985].
- 개인의 긍정적 심리 특성이 친밀한 인간관계의 형성을, 반대로 개인의 부정적 심리 특성이 부정적 인간관계를 형성하는 과정 설명
- 이후, 이 이론은 소비자 행동분야에서 밸런스 프레임워크 valence framework로 응용

■ 밸런스 프레임워크 Perters & Tarpey (1975)

- 인지된 혜택/위험 perceived benefit/risk을 개인의 의사결정의 두 가지 기본 요인으로 전체
- 인지된 혜택 측면은 개인이 긍정적 측면을 최대화하기 위해, 반면, 인지된 위험 측면은 개인이 예상되는 부정적인 영향을 최소화하기 위해 동기를 부여 받는다고 가정 [Li et al., 2018].
- 밸런스 프레임워크는 개인이 부정적 및 긍정적 효과로 인한 순가치 net value를 최대화하기 위해 의사 결정 수행 [Kim et al., 2008].

밸런스 프레임워크

- 건강 기능성 게임 수용 연구에서 밸런스 프레임워크는 적절한 이론적 토대 제시
 - 환자들(또는 일반인들 조차)은 기능성 게임을 수용하는 의사결정에 있어 발생 가능한 위험과 혜택을 동시에 고려하기 때문 [Mou et al., 2016].
 - 수용자들은 건강 기능성 게임 사용에 대한 비용과 부작용 등 예상되는 부정적 결과를 최소화하고 잠재적인 혜택을 극대화하기 위해 고민
 - 밸런스 프레임워크는 긍정적인 측면과 부정적인 측면 모두를 포함할 수 있는 고수준의 인지적 근거를 제공하여 건강 기능성 게임의 수용 결정에 영향을 미칠 수 있는 요인을 고려하는 데 적절
 - 건강 기능성 게임을 사용하기로 결정할 때, 인식된 위험으로 인한 부정적 가치와 인지된 혜택으로 인한 긍정적 가치에 모두 노출됨
 - 따라서, 건강 기능성 게임 사용 의도의 억제 요인과 촉진 요인을 모두 탐구하는 데 밸런스 프레임워크는 적절

13

밸런스 프레임워크

- 밸런스 프레임워크는 건강 기능성 게임과 같이 혜택과 위험이 동시에 작용하는 상황에서 개인 의사결정 행동을 설명하는 강력한 도구 [Gao & Waechter, 2017; Lin et al., 2014]
 - Lu et al.(2011)은 밸런스 프레임워크를 적용하여 소비자의 모바일 결제 서비스 채택 의도에 대한 인식 비용과 인식 위험의 부정적인 가치, 상대적 이점, 호환성 및 이미지의 긍정적인 가치에 대해 연구
 - Gao & Waechter(2017)는 밸런스 프레임워크를 적용하여 모바일 결제 서비스의 초기 신뢰 요인과 억제 요인에 대해 조사하였다. 이 연구에서 긍정적 가치로 시스템 품질, 정보 품질 및 서비스 품질 등을 확인하였고, 부정적 가치로 인지된 불확실성을 발견
 - Li et al.(2018)은 밸런스 프레임워크를 기반으로 모델을 개발하여 사람들이 소셜 미디어에서 건강 정보를 검색하고 공유하려는 의도를 연구
 - Xiao & Li(2019)는 긍정적인 온라인 리뷰가 가져올 수 있는 인지된 이점과 잠재적 위험을 조사하기 위해 연구 모델을 개발

14

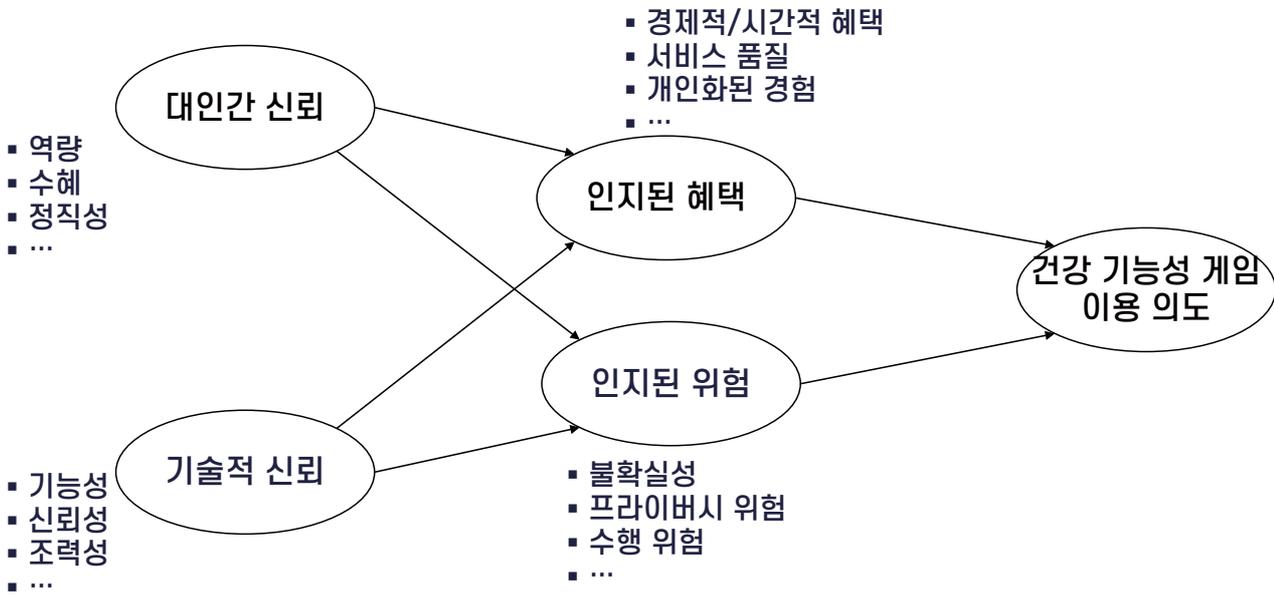


3. 연구의 방향

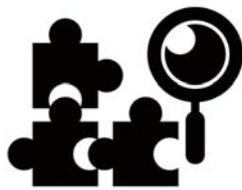
연구의 방향

- **건강 기능성 게임의 인지된 혜택과 위험을 동시 고려하여 영향 검증**
 - 밸런스 프레임워크를 기반으로 사용자의 건강 기능성 게임에 대한 수용 의사결정에 영향을 주는 촉진 요인과 저해 요인을 파악
 - 선행연구를 검토하여 건강 기능성 게임 수용자의 인지된 혜택과 인지된 위험을 측정할 수 있는 척도를 개발하고 타당화
- **연구 모델(안) 제시**
 - 밸런스 프레임워크에 대한 개념적 고찰을 통해 실증을 위한 연구 모델 제시
- **실증 연구 수행**
 - 인지된 혜택과 위험이 건강 기능성 수용에 영향을 주는 요인 검증
- **시사점 도출**

연구 모델(안)



질의 및 응답



네트워크 분석을 활용한 대학 전공 분석 : 12대 국가전략기술을 중심으로

소정은, 박병조, 김태성(충북대학교)

※ 본 논문은 저자의 요청으로 공개하지 않음.

Session B**BEXCO 제1전시장 2층 214호****B1****빅데이터와 경영인사이트****좌장 : 강주영 (아주대)**

9	개인간 연합학습 방법을 적용한 혈압 관리 시스템 : Digital Me 알고리즘을 기반으로	김수현, 정백, 이경전 (경희대)
10	AUC 최적화 학습 알고리즘의 기업 부실 예측 연구	안재현, 김명종 (부산대)
11	조건부가치측정법을 이용한 비식별화된 이미지 가치 측정	최유정, 김태성 (충북대)
12	금융 범주 불균형 최적화를 위한 AUC 부스팅 학습	권성훈, 김명종 (부산대)
13	MSA 정보화 사업의 정보시스템 감리 방안에 대한 제언	김종원 - 반짚헬 (한국과학기술정보연구원)
14	마진 분포 최적화를 통한 부실예측 모형의 예측 성과 개선	조성임, 김명종 (부산대)

개인간 연합학습 방법을 적용한 혈압 관리 시스템: Digital Me 알고리즘을 기반으로

김수현, 정백, 이경전(경희대학교)

※ 본 논문은 저자의 요청으로 공개하지 않음.

AUC 최적화 학습 알고리즘의 기업 부실 예측 연구

김명종 mjongkim@pusan.ac.kr , 안재현 bonujour8084@pusan.ac.kr

요 약

본 연구에서는 최근 활발한 연구가 진행되고 있는 기업 부실 예측에 사용할 수 있는 AUC 최적화 학습 알고리즘, AUCBoost를 제안한다. AUCBoost는 다변량 정규분포 가정을 도입하여 성과지표에 대한 직접적인 최적화가 가능하도록 수정한 알고리즘으로써 한국, 폴란드, 러시아 기업 부실 데이터에 대해 성과 개선 효과를 확인했다.

키워드

범주불균형, 기업부실예측, AUC, 성과지표 최적화, 가우시안 경사하강법

1. 서론

최근 발생한 코로나 팬데믹의 결과로 국내 부실징후 기업의 비중은 상장사 기준으로도 2009년 30.4%에서 2020년 31)9.4%로 10%p 가까이 증가하며 경제 위기에 대한 우려가 지속되고 있다. 이러한 상황 속에서 부실 기업에 대한 선제적 대응책으로 인공지능을 활용한 부실 기업 예측이 시도되고 있다. 기업 부실 예측에서 기업 부실 데이터가 가지고 있는 범주 불균형 문제는 기업 부실 예측을 위한 인공지능 개발에 있어 주요한 문제로 연구되고 있다. 선행 연구에 따르면 앙상블 학습, 데이터 분포 조작, Cost-sensitive 기법, 학습 알고리즘의 수정 등 여러 가지 방면으로 연구가 진행되고 있다. (Barboza, Kimura, Altman, 2017)

본 연구는 부실 기업 예측에 있어 알고리즘을 수정하는 방법인 AUC 최적화 학습 기법을 제안한다. AUC 최적화 학습은 다변량 정규분포 가정을 도입하여 이산확률변수로 측정되는 AUC를 미분이 가능한 연속확률함수로 치환하여 학습 알고리즘에서 성과지표에 대해 직접적인 최적화를 가능하도록 설계했다. 또한, 제안 알고리즘의 성능을 측정하기 위해 금융 환경이 다른 한국, 폴란드, 러시아 세 국가의 부실 기업 데이터를 수집하여 실험을 진행했으며, 그 결과 세 데이터 모두에서 예측 성능 개선 효과를 확인했다.

1) 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음 (IITP-2023-2020-0-01797)

2. 제안 알고리즘: AUCBoost

AdaBoost와 같은 부스팅 모형 학습 결과에 따라 AUC를 다음과 같이 도출할 수 있다. 이 때, 임의의 관측치 x 중 다수 범주 관측치를 x^- , 소수 범주 관측치를 x^+ 로 정의하며, 다수 및 소수 범주의 관측치 수는 각각 n^- 와 n^+ 로 정의한다.

$$AUC(x, w) = \frac{\sum_{i=1}^{n^+} \sum_{j=1}^{n^-} f(x_i^+, w) > f(x_j^-, w)}{n^+ n^-} \quad (1)$$

$$f(x_i^+, w) > f(x_j^-, w) = \begin{cases} 1 & \text{if } f(x_i^+, w) > f(x_j^-, w) \\ 0 & \text{Otherwise} \end{cases}$$

식 (1)에서 손실함수 $f(x_i^+, w) > f(x_j^-, w)$ 는 0 또는 1의 값을 가지는 비평활 손실함수 (non-smoothed loss function)로써 미분이 불가능하여 경사하강법과 같은 최적화 기법의 적용이 불가능하다. 식 (1)을 연속함수함수로 변환한다면 AUC의 기대값 $E(AUC(x, w))$ 을 식 (2)와 같이 계산할 수 있는데, $P_{X^+, X^-}(x^+, x^-)$ 는 다수 범주와 소수 범주의 확률변수의 결합확률로서 AUC의 기대값을 측정하기 위해서는 결합분포 $P_{X^+, X^-}(x^+, x^-)$ 의 사전적 분포에 대한 정보가 필요하다.

$$\begin{aligned} E(AUC(x, w)) &= E_{x^+, x^-}(f(x^+, w) > f(x^-, w)) \\ &= \int_{x^+} \int_{x^-} P_{X^+, X^-}(x^+, x^-) (f(x^+, w) > f(x^-, w)) dx^-, dx^+ \\ &= P(W^T X^+ > W^T X^-) = P(W^T (X^- - X^+) < 0) \quad (2) \end{aligned}$$

본 연구에서는 데이터가 충분하게 확보되는 경우 표본의 분포는 정규분포를 따른다는 중심극한정리를 활용하여 AUC의 기대값을 측정하고자 한다. 결합확률 $P_{X^+, X^-}(x^+, x^-)$ 가 다변량 정규분포를 따른다는 가정을 도입하여 다수 범주 확률변수 x^- 와 소수 범주 확률변수 x^+ 에 대한 기저분류자의 예측 결과는 식 (3)과 같이 정의할 수 있다. μ^- 와 μ^+ 는 기저분류자에서 산출된 각 범주 관측치들의 예측 결과값의 평균벡터(mean vector)이며, Σ^- 와 Σ^+ 는 각 범주 관측치들의 예측 결과값의 공분산 행렬(covariance matrix)이다.

$$E \begin{pmatrix} X^+ \\ X^- \end{pmatrix} \sim N(\mu, \Sigma) \quad \mu = \begin{pmatrix} \mu^+ \\ \mu^- \end{pmatrix} \quad \Sigma = \begin{pmatrix} \Sigma^{++} & \Sigma^{\pm} \\ \Sigma^{\mp} & \Sigma^{--} \end{pmatrix} \quad (3)$$

식 (3)에서 X^- 와 X^+ 가 다변량 정규분포를 따르게 되면 X^- 와 X^+ 의 주변확률변수도 다음과 같은 특성을 가진 다변량 정규분포로 정의된다.

$$\begin{aligned} X^- &\sim N(\mu^-, \Sigma^-) \rightarrow W^T X^- \sim N(\omega^T \mu^-, \omega^T \Sigma^- \omega) \sim N(\mu_Z^-, \sigma_Z^2) \\ X^+ &\sim N(\mu^+, \Sigma^+) \rightarrow W^T X^+ \sim N(\omega^T \mu^+, \omega^T \Sigma^+ \omega) \sim N(\mu_Z^+, \sigma_Z^2) \end{aligned} \quad (4)$$

여기에서 확률변수 X^- 와 X^+ 의 선형함수인 $Z = w^T(X^- - X^+) \sim N(\mu^Z, \sigma^Z)$ 이며 $\mu^Z = w^T(\mu^- - \mu^+)$ 이고 $\sigma^Z = \sqrt{w^T(\Sigma^- + \Sigma^+ - \Sigma^\mp - \Sigma^\pm)}$ 이다.

표준정규분포의 CDF(Φ)를 이용하여 식 (2)의 $AUC(x, w)$ 를 정의하게 되면 식 (5)와 같다.

$$AUC(x, w) = \Phi\left(\frac{\mu_Z}{\sigma_Z}\right) = -\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\mu_Z}{\sigma_Z}\right)^2\right) \quad (5)$$

식 (5)의 1차 미분함수인 $\nabla AUC(x, w)$ 는 식 (6)과 같이 정의된다. 여기에서 $\hat{\mu} = \mu^- - \mu^+$ 이며 $\hat{\Sigma} = \Sigma^- + \Sigma^+ - \Sigma^\mp - \Sigma^\pm$ 이다.

$$\nabla AUC(x, w) = -\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\mu_Z}{\sigma_Z}\right)^2\right) \left(\frac{\sigma_Z \hat{\mu} - \left(\frac{\mu_Z}{\sigma_Z}\right) \hat{\Sigma} w}{\sigma_Z^2}\right) \quad (6)$$

식 (6)에 대하여 가우시안 경사하강법을 적용하여 결합 가중치 집합(Δw_k)이 탐색되며 학습률(learning rate)와 결합하여 $w_k^{New} = w_k^{Old} - \beta \cdot \Delta w_k$ 와 같이 새로운 결합가중치(w_k^{New})를 생성하게 된다. 결합가중치는 정규화 가정($w_k^{New} / \sum_{k=1}^k w_k^{New}$)을 거쳐 후행 학습의 가중치로 활용된다. 가중치 탐색작업은 종료조건에 도달할 때까지 반복적으로 수행되며, 최종적으로 $AUC(x, w)$ 를 극대화하는 최적의 결합가중치(w^*)가 탐색된다.

3. 연구 결과

본 연구에서는 AUC 최적화 학습 알고리즘의 성과를 기존의 부스팅 기법만을 사용했을 때와 비교했다. 벤치마킹 모형으로는 AdaBoost, GBM, XGB를 사용했으며 기저 분류자는 최대 깊이가 3인 의사결정나무로 동일하게 설정했으며 폴란드, 러시아 기업 부실 데이터는 선행 연구가 이루어진 뒤 공개된 데이터를 각각 UCI Machine Learning Repository, Kaggle에서 수집했고 한국 기업 부실 데이터는 TS2000을 이용해 재무비율을 추출한 뒤 전처리 과정을 거쳐 사용했다(Zieba, Tomczak, Tomczak, 2016; Yuri, Nikita, 2021; Yun, Kim, 2021). 사용한 데이터의 간략한 현황을 <표 1>에 정리해두었다.

<표 1> 연구 데이터 개요

데이터	입력 변수의 수	소수 범주	다수 범주	불균형 비율
한국 기업 부도	7	500	10,000	1:20
폴란드 기업 부도	64	410	5,500	1:13
러시아 기업 부도	55	456	2,001	1:4

10-fold 교차검증 3번에 대한 성과를 분석한 주요 결과는 다음과 같다. 3개 데이터 모두에서 AUCBoost를 사용했을 때 산술평균 정확도(ACC)가 소폭 하락하는 결과가 나타났지만, AUC와 기하평균 정확도(GM)은 큰 폭으로 증가했다. 이는 모형의 분류 경계선이 다수 범주의 방향으로 이동한 결과로 해석할 수 있다. 즉, <표 2>의 분석 결과는 본 연구에서 제안한 AUCBoost가 범주 불균형이 존재하는 데이터에서 양 범주를 고루 반영하는 학습을 했으며 데이터와 관계 없이 강건한 학습 결과를 보여준다는 것을 실증한다.

<표 2> 모형 성과 분석 결과

Panel A. 한국 기업 부도 예측 (IR 1:20)

Performance Metrics	Training Set				Test Set			
	AdaBoost	GBM	XGB	AUCBoost	AdaBoost	GBM	XGB	AUCBoost
SPE	0.99	0.99	0.99	0.80	0.99	0.99	0.99	0.78
SEN	0.05	0.07	0.07	0.77	0.04	0.05	0.06	0.75
ACC	0.95	0.96	0.96	0.80	0.95	0.95	0.95	0.78
AUC	0.52	0.53	0.53	0.78	0.52	0.53	0.53	0.76
GM	0.22	0.26	0.26	0.78	0.18	0.21	0.22	0.76

Panel B. 폴란드 기업 부도 예측 (IR 1:13)

Performance Metrics	Training Set				Test Set			
	AdaBoost	GBM	XGB	AUCBoost	AdaBoost	GBM	XGB	AUCBoost
SPE	0.99	0.99	0.99	0.89	0.99	0.99	0.99	0.84
SEN	0.48	0.32	0.44	0.88	0.42	0.29	0.39	0.82
ACC	0.96	0.95	0.96	0.89	0.95	0.95	0.95	0.83
AUC	0.73	0.66	0.72	0.89	0.70	0.64	0.69	0.83
GM	0.69	0.56	0.66	0.89	0.64	0.53	0.62	0.83

Panel C. 러시아 기업 부도 예측 (IR 1:4)

Performance Metrics	Training Set				Test Set			
	AdaBoost	GBM	XGB	AUCBoost	AdaBoost	GBM	XGB	AUCBoost
SPE	0.96	0.99	0.98	0.81	0.95	0.98	0.97	0.75
SEN	0.49	0.33	0.46	0.81	0.43	0.30	0.40	0.73
ACC	0.88	0.86	0.88	0.81	0.85	0.85	0.86	0.75
AUC	0.73	0.66	0.72	0.81	0.69	0.64	0.68	0.74
GM	0.69	0.57	0.67	0.81	0.63	0.54	0.62	0.74

참고문헌

- Barboza, F., Kimura, H. and Altman, E. (2017). Machine learning models and bankruptcy prediction, *Expert Systems with Applications*, 83, pp.405-417. DOI: <https://doi.org/10.1016/j.eswa.2017.04.006>
- Yun. W-S, Kim. M-J (2021), AUROC-based Ensemble Model for Bankruptcy Prediction, *Journal of SME Finance*, 41(3), pp.41-60. DOI: doi:10.33219/jsmef.2021.41.3.002
- Yuri Zelenkov, Nikita Volodarskiy (2021), Bankruptcy prediction on the base of the unbalanced data using multi-objective selection of classifiers, *Expert Systems with Applications*, 185. DOI : ISSN 0957-4174, <https://doi-org-ssl.eproxy.pusan.ac.kr/10.1016/j.eswa.2021.115559>.
- Zieba, M. and Tomczak, S. K., Tomczak, J. M. (2016), Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction, *Expert Systems with Applications*, 58, pp.93-101 DOI: <https://doi.org/10.1016/j.eswa.2016.04.001>

조건부가치측정법을 이용한 비식별화된 이미지 가치 측정

최유정, 김태성(충북대학교)

※ 본 논문은 저자의 요청으로 공개하지 않음.

금융 범주 불균형 최적화를 위한 AUC 부스팅 학습

김명종*, 권성훈**

요약

범주 불균형 문제는 분류 및 예측 성과를 저하시켜 부스팅 모형의 성과 개선을 위해 해결해야 할 문제이다. 본 연구에서는 금융 분야의 범주 불균형 문제에 적용된 부스팅 계열 알고리즘의 성과 개선을 위하여 성과지표에 대한 직접적인 최적화 기법을 도입한 AUCBoost(Area Under Curve-based Boosting) 기법을 제안한다. 본 연구에서는 기업 부도, 카드 연체 및 카드 사기와 같은 대표적인 범주 불균형 문제를 대상으로 AUCBoost의 성과를 측정하였다. 성과의 비교분석을 위해 벤치마킹 모형으로 부스팅 학습 기법인 AdaBoost, GBM, XGBoost 알고리즘을 채택하였으며 30회 교차 검증을 통해 실험을 진행하여 그 결과는 다음과 같다. 첫째, 다수 범주에 초점을 맞추어 불균형 문제에서 소수 범주의 민감도를 고려하지 못하는 기존의 알고리즘과 달리 AUCBoost는 다수 범주의 특이도와 소수 범주의 민감도를 동시에 고려하는 균형 잡힌 학습을 통해 범주 불균형 문제에 유의한 성과를 보여줌을 확인하였다. 둘째, 기존 부스팅 모형과 비교하여 AUCBoost는 AUC 및 GM 등의 성과지표에서도 향상된 성과를 보여주고 있으며, 부스팅 학습의 성과지표에 관한 직접적인 최적화 적용 가능성을 시사한다.

키워드 : 범주 불균형, AUC, AUCBoost, 성과지표 최적화, 가우시안 경사하강법, 기업부도, 카드사기, 카드연체

I. 서론

4차 산업혁명과 AI 기술의 발전이 가속됨에 따라 다양한 산업과 기업들의 기술환경이 변화하고 있다. 인공지능의 도입과 디지털 전환(digital transformation)이 이루어지며 소비자의 행동과 선호도가 다각화됨에 따라 경영 패러다임이 변화하고 있다. 그러나, 기계학습의 금융 분야 도입에 앞서 부도예측, 카드사기, 대출연체 등의 문제에서 관측되는 범주 불균형 문제는 분류 및 예측모형의 성과 개선을 위해 필수적으로 고려되어야 할 문제이다(Galar et

* 소속 : 부산대학교 경영대학 교수

email : mjongkim@pusan.ac.kr

** 소속 : 부산대학교 경영학과 학사과정

email : rnstjdgns216@pusan.ac.kr

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음(IITP-2023-2020-0-01797)

al., 2012).

본 연구는 다음과 같은 수정 알고리즘을 도입하여 부스팅 모형의 성과개선을 제시한다. 첫째, 분류 모형의 목적함수를 손실함수 극소화에서 AUC 극대화로 대체하여 범주 불균형 문제에 대한 성과지표의 유효성을 확보하였다. 둘째, 이산확률함수로 측정된 AUC에 대하여 다변량 정규분포 가정(the assumption of multivariate normal distribution)을 도입하여 미분 가능한 정규분포함수로 치환하였다. 셋째, 치환된 AUC에 대하여 확률개념 기반의 가우시안 경사하강법(Gaussian gradient descent method)을 적용하여 성과지표에 대한 직접적인 최적화를 수행하였다.

II. 제안 알고리즘: AUCBoost

본 연구는 범주 불균형 문제의 해결을 위해 AUCBoost를 제안한다. 부스팅 알고리즘의 수정으로 AUCBoost가 불균형 문제를 해결하는 과정은 다음과 같다. 부스팅 모형을 통해 다수 범주($Y=-1$)와 소수 범주($Y=+1$)로 이루어진 이범주 분류문제(binary classification problem)의 학습결과를 도출한다. 그 결과를 통해 도출된 AUC는 다음과 같다. 임의의 관측치 x 중 다수 범주 관측치를 x^- , 소수 범주 관측치를 x^+ 로 정의, 다수 및 소수 범주의 관측치 수는 각각 n^- 와 n^+ 로 정의한다.

$$AUC(x, w) = \frac{\sum_{i=1}^{n^+} \sum_{j=1}^{n^-} f(x_i^+, w) > f(x_j^-, w)}{n^+ n^-} \quad (1)$$

$$f(x_i^+, w) > f(x_j^-, w) = \begin{cases} 1 & \text{if } f(x_i^+, w) > f(x_j^-, w) \\ 0 & \text{Otherwise} \end{cases}$$

단, x^- : 다수 범주 관측치

x^+ : 소수 범주 관측치

n^- : 다수 범주 관측치 수

n^+ : 소수 범주 관측치 수

식 (1)에서 손실함수 $f(x_i^+, w) > f(x_j^-, w)$ 는 0 또는 1의 값을 가지는 비평활 손실함수(non-smoothed loss function)로써 미분이 불가능하여 경사하강법 등의 최적화 기법의 적용이 불가능하다. 이를 대체하는 RMSE(root mean squared error), 로그손실(log loss), 힙지 손실(hinge loss)등의 손실함수는 범주 균형을 가정하고 있어 불균형 문제의 적용에 한계가 있다.

식 (1)을 연속확률함수로 변환하여 AUC의 기대값 $E(AUC(x, w))$ 을 식 (2)와 같이 계산할 수 있다. $P_{X^+, X^-}(x^+, x^-)$ 는 다수 범주와 소수 범주의 확률변수의 결합확률로서 AUC의 기대값

을 측정하기 위해 결합분포 $P_{X^+, X^-}(x^+, x^-)$ 의 사전적 분포에 대한 정보가 필요하다.

$$\begin{aligned} E(AUC(x, w)) &= E_{x^+, x^-} (f(x^+, w) > f(x^-, w)) & (2) \\ &= \int_{x^+} \int_{x^-} P_{X^+, X^-}(x^+, x^-) (f(x^+, w) > f(x^-, w)) dx^-, dx^+ \\ &= P(W^T X^+ > W^T X^-) = P(W^T (X^- - X^+) < 0) \end{aligned}$$

본 연구에서는 데이터가 충분히 확보되는 경우 표본의 분포는 정규분포를 따른다는 중심극한정리를 활용하여 AUC의 기대값을 측정하고자 한다. 결합확률 $P_{X^+, X^-}(x^+, x^-)$ 가 다변량 정규분포를 따른다는 가정을 도입하여 다수 범주 확률변수 X^- 와 소수 범주 확률변수 X^+ 에 대한 기저분류자의 예측 결과는 식 (3)과 같이 정의할 수 있다. μ^- 와 μ^+ 는 기저분류자에서 산출된 각 범주 관측치들의 예측 결과값의 평균벡터(mean vector)이며, Σ^- 와 Σ^+ 는 각 범주 관측치들의 예측 결과값의 공분산 행렬(covariance matrix)이다.

$$E \begin{pmatrix} X^+ \\ X^- \end{pmatrix} \sim N(\mu, \Sigma) \quad \mu = \begin{pmatrix} \mu^+ \\ \mu^- \end{pmatrix} \quad \Sigma = \begin{pmatrix} \Sigma^{++} & \Sigma^{\pm} \\ \Sigma^{\mp} & \Sigma^{--} \end{pmatrix} \quad (3)$$

식 (3)에서 X^- 와 X^+ 가 다변량 정규분포를 따르게 되면 X^- 와 X^+ 의 주변확률변수도 다음과 같은 특성을 가진 다변량 정규분포로 정의된다.

$$\begin{aligned} X^- &\sim N(\mu^-, \Sigma^-) \rightarrow W^T X^- \sim N(w^T \mu^-, w^T \Sigma^- w) \sim N(\mu_Z^-, \sigma_Z^2) & (4) \\ X^+ &\sim N(\mu^+, \Sigma^+) \rightarrow W^T X^+ \sim N(w^T \mu^+, w^T \Sigma^+ w) \sim N(\mu_Z^+, \sigma_Z^2) \end{aligned}$$

여기에서 확률변수 X^- 와 X^+ 의 선형함수인 $Z = w^T (X^- - X^+) \sim N(\mu^Z, \sigma^Z)$ 이며 $\mu^Z = w^T (\mu^- - \mu^+)$ 이고 $\sigma^Z = \sqrt{w^T (\Sigma^- + \Sigma^+ - \Sigma^{\mp} - \Sigma^{\pm})}$ 이다.

표준정규분포의 CDF(Φ)를 이용하여 식 (2)의 $AUC(x, w)$ 를 정의하게 되면 식 (5)와 같다.

$$AUC(x, w) = \Phi\left(\frac{\mu_Z}{\sigma_Z}\right) = -\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\mu_Z}{\sigma_Z}\right)^2\right) \quad (5)$$

식 (5)의 1차 미분함수인 $\nabla AUC(x, w)$ 는 식 (6)과 같이 정의된다. 여기에서 $\hat{\mu} = \mu^- - \mu^+$ 이며

$$\hat{\Sigma} = \Sigma^- + \Sigma^+ - \Sigma^{\mp} - \Sigma^{\pm} \text{이다. } \nabla AUC(x, w) = -\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\mu_Z}{\sigma_Z}\right)^2\right) \left(\frac{\sigma_Z \hat{\mu} - \left(\frac{\mu_Z}{\sigma_Z}\right) \hat{\Sigma} w}{\sigma_Z^2}\right) \quad (6)$$

식 (6)에 대하여 가우시안 경사하강법을 적용하여 결합 가중치 집합(Δw_k)이 탐색되며 학습률(learning rate)과 결합하여 $w_k^{\text{new}} = w_k^{\text{old}} - \beta \cdot \Delta w_k$ 와 같이 새로운 결합가중치(w_k^{new})를 생성하게 된다. 결합가중치는 정규화 가정을 거쳐 후행 학습의 가중치로 활용된다. 가중치 탐색작업은 종료조건에 도달할 때까지 반복적으로 수행되며, 최종적으로 $AUC(x, w)$ 를 극대화하는 최적의 결합가중치가 탐색된다.

III. 자료 및 모형설계

본 연구는 금융 분야의 범주 불균형 문제 해결을 위해 기업부도, 카드연체, 카드사기에 대한 데이터를 수집하였으며, 실증분석을 위한 분석모형은 다음과 같다. 데이터 수집을 통해 통합 DB를 구축, 각 데이터 별 부스팅 모형을 구축하여 그 성과를 비교하고자 한다. AUCBoost와 함께 부스팅 알고리즘으로 AdaBoost, GBM, XGBoost를 비교 모형으로 채택하였으며, 각 모형의 부스팅 기저분류자로 의사결정나무를 선정하였다. 각 기저분류자의 최대 depth는 2, 기저분류자의 개수는 20개로 설정하였다. 이후 교차타당성 분석을 위해 10-fold validation을 3회 실시하여 총 30회의 교차검증을 실시하였다.

IV. 결론

Table 4 수집된 데이터에 대한 부스팅 알고리즘과 AUCBoost 모형의 성과를 검증한 결과를 제시한다. 총 30회의 교차검증을 통해 특이도(SPE), 민감도(SEN), 정확도(ACC), AUC, GM의 평균값을 성과지표로 산출하였다.

Table 4. Performance Comparison of Bankruptcy, Card Fraud, and Card Insolvency Data

Panel A. Bankruptcy (IR 20:1)									
Performance Metrics	Training Set				Test Set				
	AdaBoost	GBM	XGB	AUCBoost	AdaBoost	GBM	XGB	AUCBoost	
SPE	0.99	0.99	0.99	0.80	0.99	0.99	0.99	0.78	
SEN	0.05	0.07	0.07	0.77	0.04	0.05	0.06	0.75	
ACC	0.95	0.96	0.96	0.80	0.95	0.95	0.95	0.78	
AUC	0.52	0.53	0.53	0.78	0.52	0.53	0.53	0.76	
GM	0.22	0.26	0.26	0.78	0.18	0.21	0.22	0.76	
Panel B. Card Fraud (IR 577.8:1)									
Performance Metrics	Training Set				Test Set				
	AdaBoost	GBM	XGB	AUCBoost	AdaBoost	GBM	XGB	AUCBoost	
SPE	0.99	0.99	0.99	0.96	0.99	0.99	0.99	0.93	
SEN	0.75	0.7	0.78	0.96	0.73	0.68	0.75	0.92	
ACC	0.99	0.99	0.99	0.96	0.99	0.99	0.99	0.93	
AUC	0.87	0.85	0.89	0.99	0.87	0.84	0.87	0.98	
GM	0.87	0.83	0.88	0.96	0.85	0.82	0.87	0.93	
Panel C. Card Insolvency (IR 3.52:1)									
Performance Metrics	Training Set				Test Set				
	AdaBoost	GBM	XGB	AUCBoost	AdaBoost	GBM	XGB	AUCBoost	
SPE	0.96	0.96	0.95	0.98	0.96	0.96	0.95	0.94	
SEN	0.34	0.33	0.36	0.98	0.34	0.32	0.35	0.93	
ACC	0.82	0.82	0.82	0.98	0.82	0.82	0.82	0.94	
AUC	0.64	0.64	0.66	0.98	0.65	0.64	0.65	0.94	
GM	0.57	0.56	0.58	0.98	0.57	0.56	0.58	0.94	

세가지 검증표본을 통해 기존의 알고리즘과 AUCBoost의 성과를 측정함에 따라, 기존의 알고리즘에서 특이도와 민감도의 차이가 극명하게 나타남과 달리 AUCBoost의 경우 특이도와 민감도가 유사한 수준에서 측정되어 소수 범주의 분석 성과 개선을 확인할 수 있다. AUC와 GM측면에서도 AUCBoost는 기존의 알고리즘과 비교하여 우수한 성과를 보이고 있다. 기업부도 데이터의 경우 기존 알고리즘의 AUC는 약 52~53%로 나타나는 반면, AUCBoost는 76%로 기존 알고리즘과 비교하여 약 43%의 성과 개선 효과를 보여주고 있다. 유사하게 카드 사기의 경우에도 기존 알고리즘의 AUC(84~87%)와 비교하여 AUCBoost는 94%로 약 10% 이상의 차이를 보여주고 있으며, 카드 연체의 경우에도 약 4~9%의 성과 차이를 보여주고 있다. GM 측면에서도 AUCBoost는 기존 알고리즘과 비교하여 우수한 분류 결과를 보여주고 있다.

금융 문제에서 AUCBoost는 AdaBoost, GBM 및 XGBoost와 달리 소수 범주의 민감도를 고려하여 실제 금융 불균형 데이터에서의 균형적인 학습을 진행할 수 있는 장점을 가져 기존 알고리즘의 성과의 제고를 시사한다. 본 연구는 성과지표에 대한 직접적인 최적화를 통해 기존의 범주 불균형 데이터에 적용되는 알고리즘의 문제의 개선에 대한 이론적 토대와 실증분석을 마련하였다. 특히 실제의 금융 문제에 직접적으로 적용될 수 있는 가능성을 시사하여 알고리즘 단계의 최적화와 기계학습 분야의 불균형 문제 해결에 새로운 연구 방향을 제시한다는 공헌점이 있다.

[참고문헌]

- [1] Japkowicz N, Stephen S. The class imbalance problem: A systematic study. *Intelligent Data Analysis*. 2002;6(5):429. doi:10.3233/IDA-2002-6504
- [2] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince and F. Herrera, "A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches," in *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 4, pp. 463-484, July 2012, doi: 10.1109/TSMCC.2011.2161285.
- [3] Kang, P. and Cho, S. (2006), EUS SVMs: Ensemble of Under-Sampled SVMs for Data Imbalance Problems, *ICONIP 2006: Neural Information Processing*, 837-846. DOI: https://doi.org/10.1007/11893028_93
- [4] Kubat, M., Holte, R. and Matwin, S.(1997) Learning when negative examples abound, *European Conference on Machine Learning 1997*, 146-153
- [5] Du, J., Vong, C. M., Pun, C. M., Wong, P. K. and Ip, W. F. (2017), Post-boosting of classification boundary for imbalanced data using geometric mean, *Neural Networks*, 96, 101-114. DOI: <https://doi.org/10.1016/j.neunet.2017.09.004>

MSA 정보화 사업의 정보시스템 감리 방안에 대한 제언
A Suggestion for the Information System Audit of MSA Informatization Project

김종원 / Kim Jong-Won, 한국과학기술정보연구원 [kju@kisti.re.kr]
박진형 / Park Jin-Hyung, 한국과학기술정보연구원 [toskr@kisti.re.kr]

초록

IT 기술 발전과 정보화 시대를 맞아 함에 시스템 및 서비스 개발 시 API, 클라우드 등 빠른 요구사항 대응을 위한 민첩성 중심의 경량화되어 쉽고 빠르게 적용 가능한 개발 기술 활용으로 변화되고 있다. 그에 따른 비즈니스 변화에 따른 빠른 배포, 출시, 업데이트를 수시로 진행하고 있으며, 이런 변화에 맞춘 MSA 적용을 활발하게 추진하고 있다. MSA의 기존 연구와 사례들은 MSA 적용·전환에 대한 다양한 이점 위주로만 논의되고 있으며, 실제 아키텍처 전환 시 고려사항과 구축 및 운영 시 적절한 개발 방법론과 문제사항들은 관과되고 있다. 다수의 공공기관의 정보화 사업은 정보시스템 감리를 의무화하고 있지만, MSA에 적절한 종합적인 점검 및 평가를 수행하기에는 다소 제한적이다. 본 논문에서는 국가 정보화 사업에서의 MSA 정보시스템 감리의 객관성과 신뢰성 향상에 도움을 주고 우선순위가 높은 영역을 집중적으로 감리를 수행하여 사업 위험을 최소화할 수 있다. 이를 통해 안정적인 서비스를 제공할 수 있으며, MSA 정보화사업에서의 정보시스템 감리의 표준화 및 체계화를 마련하고자 한다.

Keywords: MSA, Monolithic, Container, DevOps, Agile

1. 서론

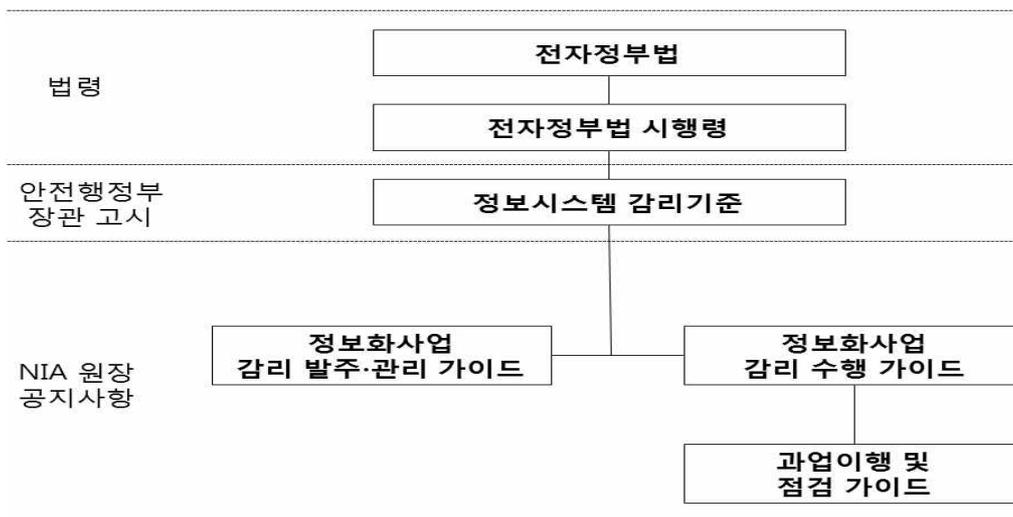
최근 기업들은 Legacy 시스템 기반 Enterprise 들의 오픈 플랫폼 전략으로 MSA 전환 필요성이 증대되고 있다. 예전처럼 대용량 서버 기반의 통합 시스템을 유지하였다면, 급속한 기술 변화, 고객 Needs 변화, 기업 비즈니스 환경 변화 등 빠른 트렌드 변화로 이제는 빠른 배포, 출시, 업데이트가 빈번하게 일어나고 있다. 이에 기업들은 서비스 구현에 있어 저렴한 클라우드 서비스나 배포가 간단한 스크립트 언어와 운영 효율화 및 자동화 등을 중점을 두고 있다. ‘Amazon Go’, 넷플릭스, 쿠팡 등 국내외 유명 기업들이 MSA 전환을 통해 디지털 전환을 가속화 하거나 서비스별 CI/CD(Continuous Integration & Continuous Deployment)를 통해 서비스를 확장하고 있다. 이를 위하여 조직 구조를 변경하거나 다양한 기술과 개발방법론도 적용한다. 하지만 MSA의 장점 이면에 전환 구축에는 많은 한계와 문제점도 존재한다. [1] 특히 일반 기업과 달리 국가 정보화사업을 추진하는 정부출연(연)이나 공공기관들의 규정, 절차, 방법론, 감리 등과 서비스적절성에 대한 논의는 부족한 실정이다.. 특히 MSA 사업에 대해 현행 정보시스템 감리에서는 감리인의 경험에 따라 감리 수준이 차이가 있고 감리의 내용과 요구 사항, 화면 설계서와 구현된 프로그램을 비교하는 수준이며, 특히 정보시스템 감리 평가가 감리원의 주관적인

판단에 의존하여 작성됨으로써 참여 감리원에 따라 상이한 평가 결과가 도출되는 현상이 발생하고 있다. 이에 감리의 객관성 및 신뢰성 향상을 위한 연구 활동이 요구되고 있으나, 객관적으로 정량적인 평가를 할 수 있는 연구가 미흡한 실정이다. 이에 본 논문에서는 국가정보화사업에서의 MSA 구축 사업시 한계점을 살펴보고, 정보시스템 감리시 MSA 구축/전화단계에서의 점검 방안들을 제안하고자 한다.

2. 정보시스템 감리

1. 정보시스템 감리의 개요

행정안전부 고시 제 2021-14호에 의거 "감리란 발주자와 사업자 등의 이해관계로부터 독립된 자가 정보시스템의 효율성을 향상시키고 안전성을 확보하기 위하여 제3자의 관점에서 정보시스템의 구축 및 운영 등에 관한 사항을 종합적으로 점검하고 문제점을 개선하는 것을 말한다" 라고 명시되어 있다. 정보시스템 감리는 1987년 국내에 처음 실시되었고, 1999년 “정보화촉진기본법” 제15조의2에 정보시스템 감리의 근거 조항이 신설되어 이를 근거로 고시된 “정보시스템 감리기준(1999년 제정, 2004년 개정)”에서는 정보시스템 감리의 절차 및 방법을 제시되고 있으며, 발생 가능한 위험을 줄이는 효과성이 인증되어 공공부문에서 그 중요성이 높게 평가되고 있다.



[그림 1] 정보시스템 감리관련 법령·고시·가이드 구성 체계[13]

정보시스템 감리에 필요한 각종 가이드는 <그림 1>와 같이 공지되고 있다. 이처럼 정보시스템 감리는 정보화 사업의 관리수준 향상, 정보시스템의 기능, 성능, 사용편의성, 정확성 등 시스템 품질 및 사용자 만족도를 높이는 효과가 인정되었으며, 정보시스템 구축의 위험요소를 사전에 최소화하여 실패를 사전에 방지하게 함으로써 재개발 또는 유지보수 비용과 노력을 절감하는 등 감리에 소요되는 비용 대비 효과가 크다는 것이 입증되었다[13]. 정보시스템 감리는 제8조 제1항 제4호에서 최종 결과물에 대한 규정하고 있으며, 이는 사업이 완전히 종료된 이후의 평가가 아니라, 사업을 진행하는 도중에 발견된 문제점 또는 위험을 사전(사업종료 이전)에 제거하기 위한 활동이며 사업유형 별, 감리시점 별로 구분하여 감리영역으로 규정하고 있다[12].

2. 정보시스템 감리 프레임워크

정보시스템 감리의 프레임워크는 [그림 2]와 같이 사업유형·감리시점, 감리영역, 감리기준·점검기준으로 크게 세 가지 축으로 구성되며, 단계별감리를 실시하는 단계 또는 시점은 정보시스템 감리 점검 프레임워크의 “사업유형·감리시점” 등을 참고하여 정하는 것을 원칙으로 하고 있다[2].



[그림 2] 정보시스템 감리 유형[14]

하지만, 정보시스템 개발 사업으로 감리기준에 따라 요구정의, 설계, 종료 3단계 감리를 해야 하는 경우에는 “정보시스템 개발 사업 3단계 감리” 부분을 함께 적용한다. 3단계 감리 유형에는 예비조사, 현장 감리, 시정조치 확인으로 구분된다. 예비조사는 감리대상사업의 현황을 파악하여 현장 감리 수행을 위한 감리계획서를 작성·제출하는 절차이다. 그리고 현장 감리는 감리계획서에 제시된 일정, 점검 항목 등에 따라 현장에서

관리를 수행한다. 그 결과를 감리수행결과보고서에 작성하여 제출하는 절차이다. 또한 시정조치 확인은 감리수행결과보고서에 제시된 개선의견이 잘 조치되었는지를 확인하고, 그 결과를 시정조치확인보고서에 작성하여 제출하는 절차로 구성된다.

3. 정보시스템 감리에서의 MSA 사업 검토 한계점

우리나라의 정보시스템 개발 방법론은 2000년대부 객체지향과 컴포넌트 기반 개발 방법론을 반영한 방법론 관리기로 발전하였다. 1990년 중반 이후에 공공부문에서는 한국형 방법론인 마르미가 개발되었다. 마르미-III는 컴포넌트기반의 시스템 개발에 필요한 기법과 작업별 산출물을 정의하고 상세한 개발 절차와 지침을 제공한다.[3][4] 대다수 공공기관의 국가 정보화 사업은 CBD에 기반한 마르미-III로 진행되었다. 최근에는 애자일 개발방법론이 대두되고 있지만, 감리 의무대상은 국가 정보화사업은 CBD 기반에 개발방법론에 따라 개발을 진행하고 산출물을 관리하여야 한다. 다수의 정보시스템은 모놀리틱(monolithic) 아키텍처와 RDB로 시스템을 구축하고 있다. 특히 국가 정보화사업중 감리 대상에 해당되는 사업은 감리를 의무적으로 시행하여야 한다. 감리가 특정 개발론에 한정되지는 않지만 CBD와 같은 전통적 개발방법론에 따른 사업에 최적화 되어 있다.[5][6] MSA를 도입하는 기관들은 빠른 비즈니스 환경 대응을 위한 단순/경량 개발환경에서 CI/CD를 목표로 조직구조도 Devops를 지향하고 있으며, 개발방법론측면에서는 애자일이 적합한 구조이다. 애자일 방법론은 개발방법론과는 [표 1]와 같은 차이가 있다. 기존 정보화 사업의 방법론에 애자일을 도입하는 다양한 연구들이 있으나, 제도 상충 등 애자일 단독으로 사업을 진행되기에는 한계가 있다.[7]

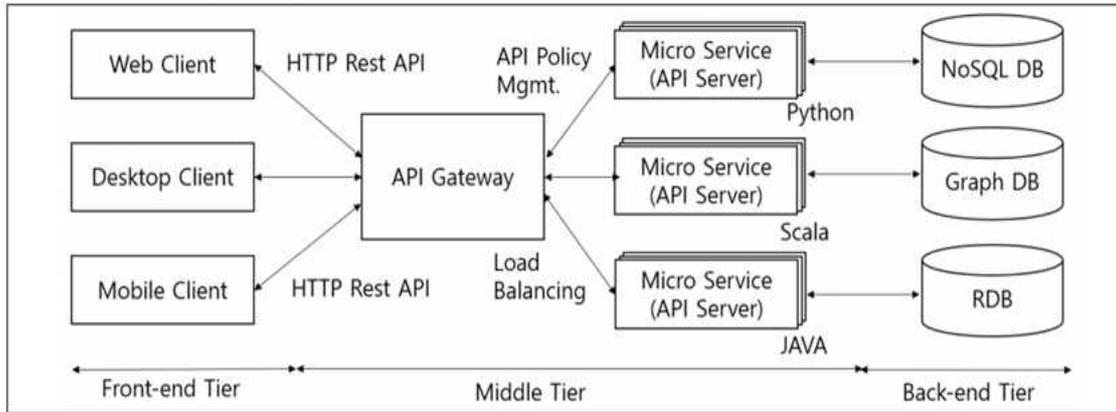
[표 1] 애자일과 전통적 개발 방법론 비교

구분	애자일 방법론	전통적인 방법론
절차	· Project backlog -> 선정 · 구현-리뷰-테스트 -> 잦은배포	· 분석-설계-구현-테스트-배포 등 순차적 하나의 Cycle 로 진행 · 잦은 배포나 반복이 어려움
방법	· 릴리즈, 리뷰	· 작업 명세서, 요구사항 추적표
산출물	· 절차보다는 사람중심 · Backlog	· 단계별 산출물, 계획서, 요구사항정의 명세서 등
관점	· 고객만족	· 산출물 품질
기법	· 칸반, XP, Scrum 등	· 객체지향, CBD 등
장점	· 민첩, 경량, 변화대응	· 적용사례 많음, 명확한 테스트 케이스
단점	· 감리 등 대응 제한	· 계획 및 절차 중심, 신속 배포 등 제한

3. 마이크로 서비스 아키텍처

3.1 MSA의 개요

MSA는 하나의 큰 어플리케이션을 아주 작은 단위로 동작하는 다수의 마이크로 서비스가 동작하는 시스템 및 소프트웨어 구성과 구성요소들간의 관계를 정의한 아키텍처이다.[8][9] MSA는 [그림 3]와 같이 개발언어 및 DB 등이 상이한 각 서비스별 API 통신을 하는 분산 구조이다. API Gateway를 통해 클라이언트와 마이크로서비스간 통신이 가능한 구성요소를 이루고 있으며, 구성요소별 주요 기능은 [표 2]와 같다.

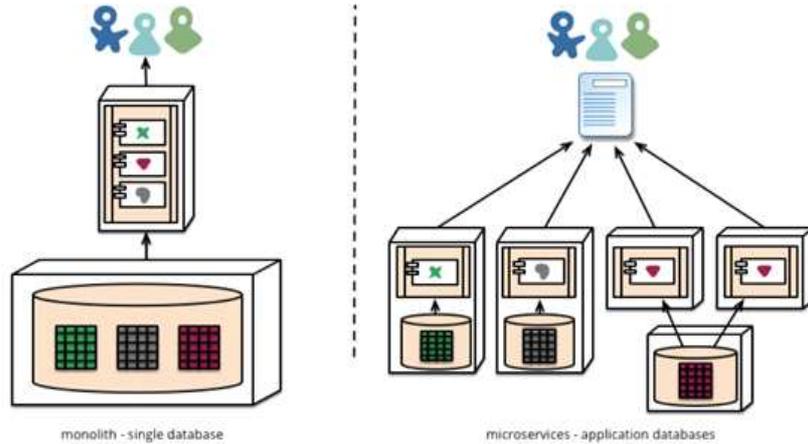


[그림 3] MSA 구조

[표 2] MSA 구성요소

구분	구성요소	설명
Fronte-end Tier	· UI/UX	· Web, Mobile, Desktop 등의 클라이언트 어플리케이션
	· Protocol	· HTTP Rest API, AMQP 등의 경량화된 프로토콜
Middle Tier	· API Gateway	· API 서비스 로드밸런싱, 인증/로깅 등의 공통 기능 제공
	API Server	· 상호 독립적으로 배포/관리 가능한 단위의 개별 서비스
Back-end Tier	NoSQL DB	· Document, Graph, Key/Value 기반의 수평적 확장 가능한 DB
	RDB	· ACID를 보장하는 관계형 데이터 모델에 기초를 둔 DB

3.2 MSA의 장단점



[그림 4] 모놀리틱과 MSA 아키텍처 차이

[그림 5]은 모놀리틱 아키텍처와 MSA를 비교한 것이다. MSA는 모놀리틱 아키텍처에 비해 독립적 서비스별 구축과 배포가 가능하며 독립적 DB 환경을 구축하여 개발 속도가 빠르며 부분적 장애처리와 다양한 언어지원 등 [표 1]과 같이 다양한 장점이 있다. 반면, 복잡한 구조로 인해 서비스가 커짐에 따라 운영·관리 측면에서 단점도 존재하며, 다른 아키텍처와는 달리 [표 2]와 같이 조직관점에서의 고려도 필요하다.[10]

[표 3] MSA의 장점

구분	주요기술	내용
시스템 확장성	· 컨테이너 기반 신속 배포/확장	· 특정서비스에 대한 확장성 용이
서비스 적시성	· CI/CD · Cross Funtional Team	· 요구사항 신속 반영, 빠른 배포
서비스 독립성	· Polyglot 구조 · Restful API	· 서비스별 개별 배포 (배포시 전체 서비스 중단 없음)
시스템 안정성	· API G/W, · Circuit Breaker	· 장애가 전체로 확장되지 않음

[표 4] MSA의 단점

구분	내용	
기술적 측면	· 장애추적 및 모니터링	· 하나의 메뉴가 다양한 서비스로 구성시 추적
	· 트랜잭션 제어	· 여러 서비스에 분산되어 있는 데이터를 하나의 트랜잭션 처리 불가
	· 성능적 오버헤드	· API 통신 시 JSON이나 XML 데이터로 전달하여 Marshring 오버헤드
조직적 측면	· 서비스간 의존성	· 관련 개발 조직간 책임 우선순위 문제
	· 테스트시 조직간 협업	· 각 조직간 테스트 환경 구축 문제로 복잡도 증가
	· 역할중심 팀모델 해체	· 조직내 거버넌스 변경 필요, 인력 자원 배분 관리방안 모색 필요

4. 정보시스템 감리 제언

4.1 국가 정보화 사업 감리에서의 MSA 적용 고려사항

국가 정보화 사업은 다수가 용역형태로 진행되며 사업기획부터 완료까지 준수해야 하는 규정과 절차에 따라야한다. 이런 제도적 제한은 간과하고 단지 시스템 관점에서만 컨설팅 및 ISP 수행결과로 MSA 전환이 언급되는 경우가 많다. 정부출연(연)의 서비스 같은 경우는 R&D정보를 수집·구축·제공 등이 많기에 데이터의 정확성이 매우 중요한 경우가 많다. 일부 애자일을 접목하고 있다고는 하지만 다수의 공공기관은 CBD 중심의 개발방법론을 따라야 하며, 산출물이 중요하고 감리를 진행해야 한다. 공공기관에서 추진하는 정보화에서 조직, MSA 이해도, 서비스 특성, 서비스 개발/운영 측면에서 MSA로의 적용은 많은 부분에서 상충되거나 제한이 되는 부분이 많다고 할 수 있다. [표 5]은 정보화 사업 수행시 MSA 전환에 대한 고려사항을 정리하였다. 단순한 차세대와 같은 MSA 전환보다는 무엇보다 서비스의 특성을 고려하여 MSA를 내외부 아키텍처를 이해하고 MSA 적용을 위한 조직구조의 변경과 애자일 방법론을 적용한 CI/CD 방안으로의 변화가 선행되어야 하겠다.

[표 5] 고려사항

구분		내용
기술적 측면	· DevOps	· 자체 개발과 운영환경이 통합된 조직으로 변경, 역량 강화 필요 - 년단위 사업 발주 방식의 변화 필요
	· Agile	· 지속, 반복, Adaptive 방법론 - 콘웨이 법칙 적용, CBD단계별 기법 적용 필요
아키텍처 측면	· Inner 아키텍처	· 관련 개발 조직간 책임 우선순위 문제
	· Outer 아키텍처	· 각 조직간 테스트 환경 구축 문제로 복잡도 증가
인프라 측면	· 쿠버네틱스	· 컨테이너 관리 및 오케스트레이션 위한 운영 능력 확보
	· 모니터링	· 장애관리, 성능관리, 네트워크 병목 관리
	· API G/W	· API G/W의 성능, 모니터링 강화
개발/배포	· 서비스 분해 효율성	· 단일 정보, 소규모 서비스 경우 모놀리식이 효율적
	· DB 모델링	· No-SQL 등은 Join 불가 및 성능저하
	· 서비스 설계	· 디자인 패턴
	· CI/CD	· 변경사항 정기적 빌드, 테스트, 릴리즈 위한 조직/절차
제도	· 감리, 보안	관련 의무 규정 준수, 산출물/증적 관리

4.2 MSA 감리의 정량적 평가 개요

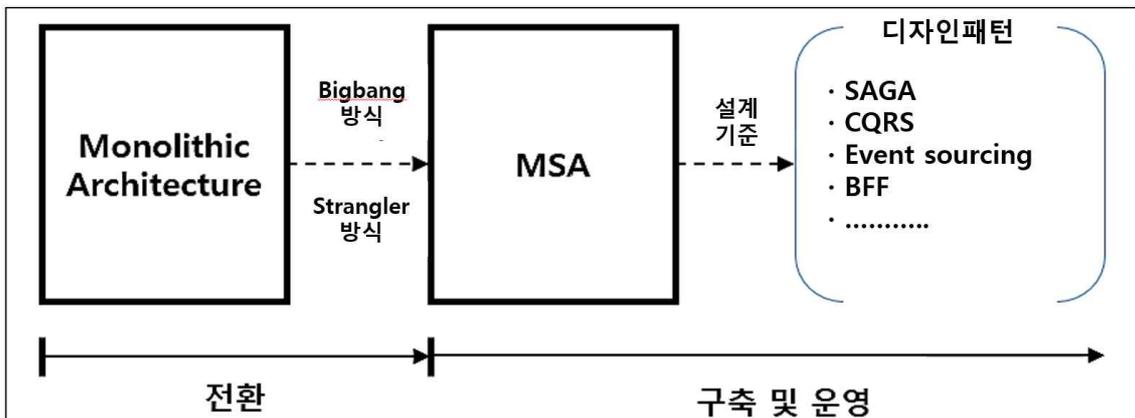
정보시스템 감리는 감리 발주기관 및 피 감리인의 이해관계로부터 독립된 자가 정보시스템의 효율성을 향상 및 안전성을 확보를 위하여 제 3자적 관점에서 정보시스템의 구축 및 운영에 관한 사항을 종합적으로 점검하고 문제점을 개선하도록 하고 정보시스템 개발 사업의 성공적인 수행을 지원하는 역할을 수행하고 있다[11]. [그림 6]과 같이 현행 정보시스템 감리 프레임워크의 시스템 개발 사업 유형에는 요구분석, 분석/설계, 구현, 시험, 전개 감리시점으로 구성되어 있다. 국가정보화사업에서의 MSA 정보시스템 감리 시스템 개발은 전환단계시 시스템 아키텍처의 감리영역과 시스템 운영 영역에서의 종합적인 연계 감리가 필요하다. 단 시스템 운영영역은 현재 운영중인 시스템의 영역이 아니므로, MSA에 대한 조직구조와 운영부분이 분석/설계단계에서 함께 이루어졌는지 확인하는 것이 중요하다.



[그림 6] MSA 감리 중점 영역

4.3 MSA 정보사업에서의 감리 평가 제언

감리는 사업자의 각 요구사항 분석이 최종 시스템 설계까지 반영되어 각 영역에서의 정의된 산출물이 잘 정리는지 확인하는 절차로 진행이 된다. 감리가 의무사항이라고 해도, MSA 사업시 시스템 전환방식이나 적용 디자인패턴에 대한 정의와 CI/CD 조직이 없다는 이유로 부적합을 할 수는 없는 것이 현실이다. 이에 현재 CBD 위주의 감리 제도와 점검 항목 유지하면서, 요구분석 단계에서의 적절성을 확인하는 항목을 추가하는 방안이 필요하다. 이를 기반으로 MSA의 감리영역을 절차화 하면 <그림7>과 같다.



[그림 7] MSA 감리 중점 영역

요구분석 단계에서는 [표 6]와 같은 감리 점검 항목 추가가 필요하다. 특히 MSA는 서비스 분해가능성과 규모확장성, 조직이 CI/CD를 고려해서 사업을 추진하고 있는지 고려되어야만 MSA 특성에 맞는 설계가 가능할 것이다. 설계영역에서는 각 디자인 패턴별 장단점을 고려하여 패턴의 적절성을 운영 부분은 감리영역의 운영 영역과 결부하여 점검을 세분화가 가능하다.

[표 6] MSA의 요구분석 영역 추가 감리 점검 항목

구분	항목	
사전 분석	· MSA 필요성	· 모토리틱 애플리케이션의 적절성/아키텍처 변경 검토
	· 분해 용이성	· 애플리케이션 내에서 비즈니스 범위 식별 여부
		· 마이크로 서비스로 충족해야 하는 요구사항 결정 여부
	· 조직구조	· 마이크로 서비스 활용 팀 재정비 여부
		· DevOps 및 CI/CD에 대한 모범사례 도입 여부
	· 설계/구현	· 디자인 패턴
· 운영	· 마이크로 서비스 오케스트레이션 구현 여부	
	· MSA 관리도구와 프로세스 구현 여부	

[표 7]와 같이 설계영역에서는 기존 감리의 점검영역처럼 요구사항과 연계하여 기술적 항목을 선정할 수 있다. MSA의 특징을 고려하여 설계에서부터 테스트는 물론 운영부분까지 고려하여 감리 영역 선정이 필요하다 하겠다.

[표 7] MSA의 설계 영역 추가 감리 점검 항목

구분	항목	설명
API 설계	· 하위 호환성	· 서비스명/버전/리소스 형태 적절성 확인
	· 독립된 버전	
REST API	· Resource	· 자원 정의
	· Verb	· 자원 행위 정의
	· Representations	· 자원 행위 내용 정의
테스트 체계	· API 명세서 준수	· swagger UI를 통한 API 테스트 확인
	· 자동화 및 모니터링	· RESTful API 문서 확인
지속적 통합 및 배포체계 설계	· Project Pipeline	· 소스코드 관리형태
	· CI Pipeline	· 조직 프로세스 수립
	· CD Pipeline	· 자동화 시각과 방안 확인
모니터링 체계 설계	· 서비스 모니터링	· 개별 서버 모니터링 방안
	· 데브옵스 모니터링	· 적용 디자인 패턴별 모니터링 적절성

5. 결론 및 향후 연구 과제

MSA는 기존 아키텍처 대비 다양한 장점기반의 구축 및 전환 사례 등이 부각되고 있다. 그에 따라 일부 공공기관등에서는 ISP 연구결과에 MSA로 전환이 필요하다고 언급되기도 한다. 하지만 국가 정보화 사업 제도하에서 MSA 서비스 특징과 조직 구조를 고려하지 않고 단순 인프라 전환이나 SW 개발 사업으로 진행하게 되면 개발에서의 어려움은 물론이고 차후 서비스 운영이나 인프라관리 등 다양한 측면에서 Risk가 증가 할 수 있다. 본 논문에서는 정보시스템 감리 시 MSA 를 조직, 아키텍처, 인프라, 개발, 제도 등 다양한 측면에서 고려사항을 도출하여, 현재 감리영역/제도/점검 항목들을 유지하면서도 MSA 사업의 효과적인 감리를 위한 추가 점검 항목에 대하여 제언하였다. MSA는 기존 아키텍처나 개발방법론과는 달리 개발-운영의 조직형태 및 방법론이 같이 고려되어야만 한다. 공공기관에서 국가 정보화 사업 추진 시 기존 서비스 구조에서 마이크로 서비스가 적절한지 판단 후, 조직-개발-운영 등 다양한 관점을 통합적으로 검토를 통해 MSA로의 전환을 검토해야 한다. 차후 MSA 사업의 특성을 고려하여 감리기법 등에 대한 연구가 필요할 것이다.

5. 참고문헌

- [1] Y.S. Kim, Understanding MSA and Using Container Technology, IITP. pp.12-25, 2019
- [2] Information System Supervision Implementation Guide, NIA, (2021)
- [3] J. J. Jahng, S. D. Kwon, S.H.Koh, ‘A Study on the History of IS Development Methodology’, Information Systems Review, pp. 211-231 (2008), Vol10. No.2.
- [4] https://etri.re.kr/45th/sub05_10.html, Jan 2, 2023
- [5] D.W. Moon, A Study on the Application Criteria of System Development Methodology, Korea Computer Institute, (1997)
- [6] Information System Supervision Implementation Guide, NIA, (2021)
- [7] D.H.KIM, “The Audit Method for Improving The Project Quality by Applying the Agile Methodology” The Graduate School of Konkuk University. Master Degree. pp.4-26. (2011)
- [8] Y.S. Kim, Understanding MSA and Using Container Technology, IITP. pp.12-25, 2019
- [9] <https://martinfowler.com/articles/microservices.html>, Jan 2, 2023
- [10] <https://martinfowler.com/bliki/MicroservicePremium.html>, Jan 2, 2023
- [11] 이기영 외, “정보시스템 감리평가 품질 향상을 위한 정량화 방법”, 한국컴퓨터정보학회논문지, 제17권 제4호 통권 제97호, 2012

마진 분포 최적화를 통한 부실예측 모형의 예측 성과 개선

조성임¹, 김명종²

요약

본 연구는 범주 불균형 환경에서 기업부실 예측모형의 성과개선을 위하여 SVM에 마진분포이론을 반영한 모형인 LDM을 적용하고 LDM을 수정한 제안모형을 제안하고자 한다. 제안모형은 LDM이 범주 불균형 문제에 보다 적합하도록 범주별 분포를 구분하여 반영한 것이다. 주요 실험 결과로 기업부실 표본에서 제안모형은 LDM과 SVM의 성과를 유의적인 수준에서 크게 개선하였다.

주요용어 : 마진분포모형, LDM, SVM, 범주 불균형, 부실예측

1. 서론

기업부실로 인하여 발생하는 손실을 최소화하고자 부실예측 모형과 관련된 많은 연구들이 진행되어 오고 있다. 통계 기법과 더불어 인공지능경망, 서포트 벡터 머신(Support Vector Machine, SVM)과 같은 다양한 기계학습이 활용되고 있으며 특히, SVM은 서포트 벡터와 분류 경계선 사이의 거리인 마진을 최대화하여 경계선을 구분하기 때문에 일반화 능력이 높아 부실예측 문제에 탁월한 성과를 보여주었다.

하지만, 범주 불균형 문제는 SVM의 분류 성과에 매우 부정적인 영향을 미치는 것으로 보고되고 있다(조성임과 김명종; 2022). 이에 대하여 cost-sensitive SVM과 같은 수정모형은 SVM의 성과를 크게 개선하였으나 마진의 크기가 서포트 벡터에 의해 제한됨에 따라 범주 불균형이 극심한 상황에서 성과향상이 제한된다(Yang et al., 2009).

이에 대하여 최근에는 서포트 벡터 대신에 전체 데이터의 마진의 분포를 반영할 경우 마진의 크기를 더욱 크게 할 수 있어 분류 모형의 일반화 성과를 더욱 높일 수 있다는 주장이 있다. 마진의 분포를 반영하는 모형은 기존의 마진 이론이 제한된 마진에 의존하여 모형의 성능을 결정하기 때문에 데이터의 수가 극히 적거나 범주 불균형과 같은 특수한 상황에서는 일반화 능력이 떨어지므로, 마진의 분포가 보다 중요하다는 주장에 근거한다(Reyzin and Schapire, 2006; Zhang and Zhou, 2013). Zhang and Zhou(2013)는 평균 마진은 최대화하는 동시에 마진의 분산은 최소화하는 마진 분포 이론인 Large margin Distribution Machine(LDM)을 제시하였다. LDM은 SVM의 수정모형으로 평균 마진을

* 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학CT연구센터육성지원사업의 연구결과로 수행되었음(ITP-2023-2020-0-01797).

¹ 46241 부산광역시 금정구 부산대학교 63번길 경영관 부산대학교 경영학과 수료후연구생. E-mail: reina337@pusan.ac.kr

² 46241 부산광역시 금정구 부산대학교 63번길 경영관 부산대학교 경영학과 교수. E-mail: mjongkim@pusan.ac.kr

최대화하여 편차를 줄이는 동시에 마진의 분산을 최소화하여 편차의 분산을 줄임으로써 분류 정확도를 향상시키는 것을 목적으로 하며 범주 불균형 상황에서도 기존의 SVM에 비하여 일반화 성과가 높은 것으로 보고하였다(Zhang and Zhou, 2013; Abe, 2017).

본 연구는 LDM이 범주 불균형 문제에 더욱 적합하도록 부실기업과 정상기업의 데이터의 분포를 고려한 알고리즘을 제안하고 SVM, LDM과 성과를 비교한다. 제안 모형은 LDM이 다수 범주의 특이도와 소수 범주의 민감도를 고려하였다는 공헌점이 있다. 본 연구의 주요 결과로 LDM과 제안모형은 SVM의 성과를 크게 개선시켰다.

본 연구는 다음과 같이 구성된다. 제 2 장은 SVM, cost-sensitive SVM, LDM, 제안모형의 알고리즘을 소개한다. 제 3 장은 연구 설계에 대하여 설명한다. 제 4 장에서 연구결과를 제시하고 제 5 장은 결론이다.

2. 알고리즘

일반적으로 알려진 SVM의 목적식과 제한식은 식 (1)과 같다(Cortes et al., 1995).

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ \text{s. t.} \quad & y_i w^T \phi(x_i) \geq 1 - \xi_i \\ & \xi_i \geq 0, \quad i = 1, \dots, n \end{aligned} \quad \text{식 (1)}$$

Cost-sensitive SVM(CSSVM)은 소수 범주의 오분류 비용을 다수 범주의 오분류 비용보다 크게 설정하며 식 (2)와 같이 정의된다(Veropoulos et al., 1999).

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + C^+ \sum_{i|y=+1}^n \xi_i + C^- \sum_{i|y=-1}^n \xi_i \\ \text{s. t.} \quad & y_i w^T \phi(x_i) \geq 1 - \xi_i \\ & \xi_i \geq 0, \quad i = 1, \dots, n \end{aligned} \quad \text{식 (2)}$$

관측치 (x_i, y_i) 과 분류 경계선 $y = w^T \phi(x)$ 의 거리(마진, γ_i), 평균 마진(γ_m)과 마진의 분산(γ_v)은 다음과 같다(Zhang and Zhou, 2013).

$$\begin{aligned} \gamma_i &= y_i w^T \phi(x_i) \\ \gamma_m &= \frac{1}{n} \sum_{i=1}^n y_i w^T \phi(x_i) \\ \gamma_v &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n (y_i w^T \phi(x_i) - y_j w^T \phi(x_j))^2 \end{aligned} \quad \text{식 (3)}$$

LDM은 SVM에 평균 마진과 마진의 분산을 적용한 것으로 식 (4)과 같이 나타낼 수 있다(Zhang and Zhou, 2013).

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + \lambda_v \gamma_v - \lambda_m \gamma_m + C \sum_{i=1}^n \xi_i \\ \text{s. t.} \quad & y_i w^T \phi(x_i) \geq 1 - \xi_i \end{aligned}$$

$$\xi_i \geq 0, \quad i = 1, \dots, n \quad \text{식 (4)}$$

여기에서 λ_v 과 λ_m 은 평균 마진과 마진의 분산에 적용되는 매개 변수이며, $\lambda_v = \lambda_m = 0$ 일 때 LDM은 SVM과 동일하다.

본 연구의 제안 모형은 LDM의 평균마진과 마진의 분산이 다수 범주와 소수 범주의 분포를 반영하도록 다음과 같이 수정한 것이다.

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + \theta_v^+ \gamma_v^+ + \theta_v^- \gamma_v^- - \theta_m^+ \gamma_m^+ - \theta_m^- \gamma_m^- + C \sum_{i=1}^n \xi_i \\ \text{s. t.} \quad & y w^T \phi(x) \geq 1 - \xi_i \\ & \xi_i \geq 0 \end{aligned} \quad \text{식 (5)}$$

여기에서 θ_m^+ 와 θ_m^- 는 평균마진 적용되는 매개 변수를 범주별로 구분하여 표현한 것이며 θ_v^+ 와 θ_v^- 는 마진의 분산에 적용되는 매개 변수를 범주별로 구분하여 표현한 것으로 이 매개변수들이 모두 0일 경우 제안모형은 SVM과 동일하다.

3. 연구설계

3.1. 표본수집과 변수선정

본 연구의 표본은 국내 시중은행으로부터 수령한 데이터를 기준으로 수집가능한 2016부터 2019년까지 국내 비금융업 기업을 대상으로 하였다. 부실의 정의는 원리금 상환기일 기준 연체 90이상인 기업, 청산 폐업 등 법적 부실기업, 채무재조정 대상 기업으로 500개를 부실 기업으로 선정하였으며, 부실이 아닌 기업 중 10,000개 정상 기업으로 선정하여 총 10,500개의 표본을 구성하였다.

본 연구는 설명변수를 선정하기 위하여 부실예측 연구에서 주로 사용되는 재무비율 30개에 대하여 7개 재무비율 군(수익성, 부채상환능력, 레버지리, 자본구조, 유동성, 활동성, 규모)으로 분류하였으며, 최종 입력변수는 각 재무비율 군별로 AUC가 가장 높은 재무비율을 선정하였다(Kim et al., 2015).

3.2. 성과지표

<표 1> 정오분류표와 다양한 성과 지표

		Prediction	
		Positive	Negative
Real	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

$$\text{Sensitivity (SEN)} = TP / (TP + FN)$$

$$\text{Specificity (SPE)} = TN / (FP + TN)$$

$$\text{Arithmetic mean accuracy (ACC)} = (TP + TN) / (TP + FN + FP + TN)$$

$$\text{Geometric Mean accuracy (GM)} = \sqrt{\text{Sensitivity} \times \text{Specificity}}$$

<표 1>은 이분류 모형의 분류 결과를 빈도로 표시한 것이다. TN 와 TP 는 각각 다수 범주와 소수 범주의 정분류된 관측치이고 FN 와 FP 은 각각 다수 범주와 소수 범주의 오분류된 관측치의 수이다. 이를 기준으로 소수 범주의 정확도(민감도)(Sensitivity, SEN), 다수 범주의 정확도(특이도)(Specificity, SPE)를 계산할 수 있다.

일반적인 정확도(Arithmetic mean accuracy, ACC)는 범주 불균형 문제에서 다수 범주의 정확도가 높아 편향되는 문제가 있다. 이를 해결하기 위하여 본 연구는 다수 범주 정확도와 소수 범주 정확도를 모두 반영하는 기하평균정확도(Geometric Mean accuracy, GM)를 모형의 일반화 성과를 측정하는 지표로 사용하고자 한다.

4. 연구결과

본 연구는 각 모형이 모형별로 최고의 성과를 낼 수 있도록 매개변수의 조합을 그리드 탐색 기법으로 탐색하였다. 손실계수는 [1, 10, 100, 200, 500, 800] 중에서 선택되었으며 RBF커널의 분산은 [2^{-2} , 2^{-1} , 1, 2^1 , 2^2] 중에서 선택되었다. LDM의 λ_v 와 λ_m 와 제안모형의 θ_v 와 θ_m 는 [0.1, 1, 10, 100, 200, 500, 800] 중에서 선택되었다.

<표 2> 부실표본의 모형별 분류성과 평가 결과

모형	정확도 (% , %p)				
	ACC	GM	SEN	SPE	SPE- SEN
SVM	95.3	4.2*	0.6	100	99.4
CSSVM	47.5	65.0*	93.6	45.2	48.4
LDM	61.3	71.8*	86.0	60.0	26
제안모형	68.5	74.7	82.4	67.8	14.6

*는 5% 수준에서 유의

<표 2>는 표본에 대하여 SVM, CSSVM, LDM 및 제안모형을 적용하여 분류한 결과로, 10-fold 교차타당성 분석을 3회 반복하는 과정을 거쳐 산출한 정확도의 평균값을 제시한 것으로 정확도 별로 가장 높은 값을 굵게 표시하였다. 주요 분석결과를 요약하면 다음과 같다. 첫째, SVM은 다른 세가지 모형에 비하여 다수범주의 특이도(100%)가 가장 높았으며 CSSVM은 소수범주의 민감도(93.6%)가 가장 높게 나타났다. 그 결과 두 모형에서 다수 범주의 특이도와 소수 범주의 민감도의 차이가 SVM은 99.4%p, CSSVM은 48.4%p로 크게 나타났다. 이와 같은 결과로 SVM과 CSSVM은 다수 범주의 특이도와 소수 범주의 민감도를 균형 있게 반영하지 못하고 하나의 범주에 지나치게 의존하여 일반화성과를 나타내는 기하평균 정확도(GM)가 낮게 나타났다. 둘째, LDM과 제안모형에서 다수 범주의 특이도와 소수 범주의 민감도의 차이를 비교하면 LDM은 26%p, 제안모형은 14.6%p로 SVM과 CSSVM에 비하여 큰 폭으로 낮아졌다. 그 결과 제안모형과 LDM의 기하평균정확도는 SVM과

CSSVM의 기하평균정확도를 크게 개선하였다. 또한 제안모형의 기하평균정확도는 4가지 모형 중에서 가장 높은 값을 보인다. 이러한 결과는 마진의 분포를 반영한 모형이 서포트 벡터에 제약받지 않고 서포트 벡터를 넘어서는 범위에까지 마진의 크기를 크게 하고, 특이도와 민감도의 차이를 최소화하여 일반화 능력을 개선하는 최적의 조건을 찾은 것으로 해석될 수 있다. 또한 제안모형은 SVM과 LDM에 비하여 다수범주와 소수범주의 분포를 각각 반영하여 특이도와 민감도가 가장 근접한 지점을 찾은 것으로 해석된다.

5. 결론

본 연구는 SVM이 범주 불균형 문제에서 일반화성능을 개선하기 위하여 기존의 마진이론 대신에 마진 분포 이론을 고려하였다. 또한 기존 연구에서 제시한 SVM에 마진의 분포를 반영한 LDM이 범주 불균형 문제에 특화될 수 있도록 평균마진과 마진의 분산에 다수 범주와 소수 범주의 분포를 반영한 제안모형을 제안하였다. 분석 결과, 제안모형은 SVM과 LDM의 일반화 성과를 유의적으로 크게 개선하였다. 본 연구는 범주 불균형 문제를 해결하기 위하여 일반적으로 사용되는 데이터 샘플링 기법이 데이터의 원형이 가지는 정보를 훼손하는 점을 해결하고자 알고리즘 수정 기법을 적용하였고, 제안모형이 기존에 제시된 수정 기법에 비하여 우수한 분류 성과를 확인하였다는 점에서 공헌점이 있다.

[References]

- 조성임과 김명중(2022). “비대칭 마진 SVM 최적화 모델을 이용한 기업부실 예측모형의 범주 불균형 문제 해결”. *Information Systems Review*, 24(4). 23-40.
- Breiman, L. (1999). “Prediction Games and Arcing Algorithms”. *Neural Computation*, 11(7), 1493-1517.
- Cortes, C., Vapnik, V., and Saitta, L. (1995). “Support-Vector Networks Editor”. *In Machine Learning*, 20. Kluwer Academic Publishers.
- Reyzin, L., and Schapire, R. E. (2006). “How boosting the margin can also boost classifier complexity”. *Proceedings of the 23rd International Conference on Machine Learning - ICML '06*, 753-760.
- Veropoulos, K., Campbell, C., and Cristianini, N. (1999). “Controlling the Sensitivity of Support Vector Machines”. 1-6.
- Yang, C. Y., Yang, J. S., and Wang, J. J. (2009). “Margin calibration in SVM class-imbalanced learning”. *Neurocomputing*, 73(1-3), 397-411.
- Zhang, T., and Zhou, Z.-H. (2013). “Large Margin Distribution Machine”. *KDD '14: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 313-322.