

A man in a dark suit and tie is pointing with a pen at a futuristic, glowing blue digital interface. The interface features the large text 'ESG' and a central circular graphic with a globe icon. To the right of the circle are six circular icons arranged in two rows of three, representing various ESG themes like environment, social, and governance. On the left, the words 'ENVIRONMENT', 'SOCIAL', and 'GOVERNANCE' are listed vertically with corresponding lines pointing to the central graphic.

ESG

ESG 경영 변화 추적 및 전략적 대응을 위한 ESG 대시보드 제작

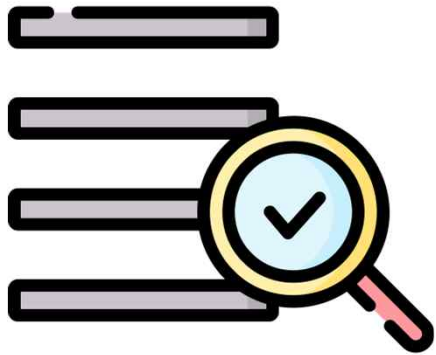
ENVIRONMENT

SOCIAL

GOVERNANCE

00

목차



01

프로젝트 소개

프로젝트 정의
프로젝트 배경
프로젝트 목표

02

프로젝트 정보

프로젝트 기간
서비스 플로우
프로젝트 진행 프로세스
개발 환경 및 적용 기술 스택
주요 기능 구현 상세
성과 및 결과

03

회고

Issue
후기

04

참고자료

리서칭 자료

01

프로젝트 소개

프로젝트 정의



웹 크롤링 기반 비정형 데이터 분석 모델을 통한
공시 기업별 ESG 적용 중요도 Score의 AI 기반 자동 산출 시스템 개발

01

프로젝트 소개

프로젝트 배경

ESG의 필요성

ESG란? (출처: 한국거래소)

Environment(환경), Social(사회), Government(지배구조)의 약자
기업 경영에서 지속 가능성을 달성하기 위한 3가지 핵심 요소



ESG 필요성 대두

전통적 재무보고 방식의 한계를 극복하고 보완할 수 있는 새로운 평가 지표 필요 (출처: 자본시장연구원, 2021)

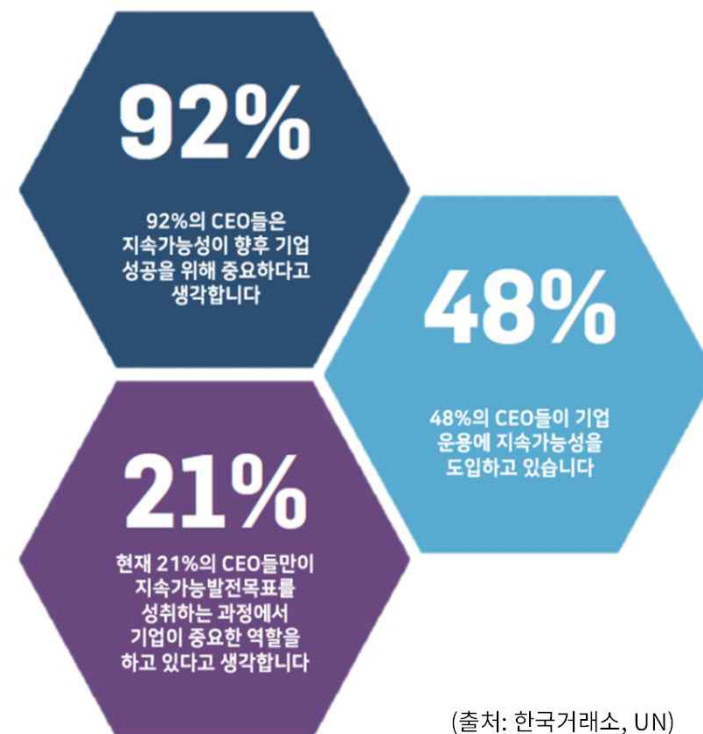
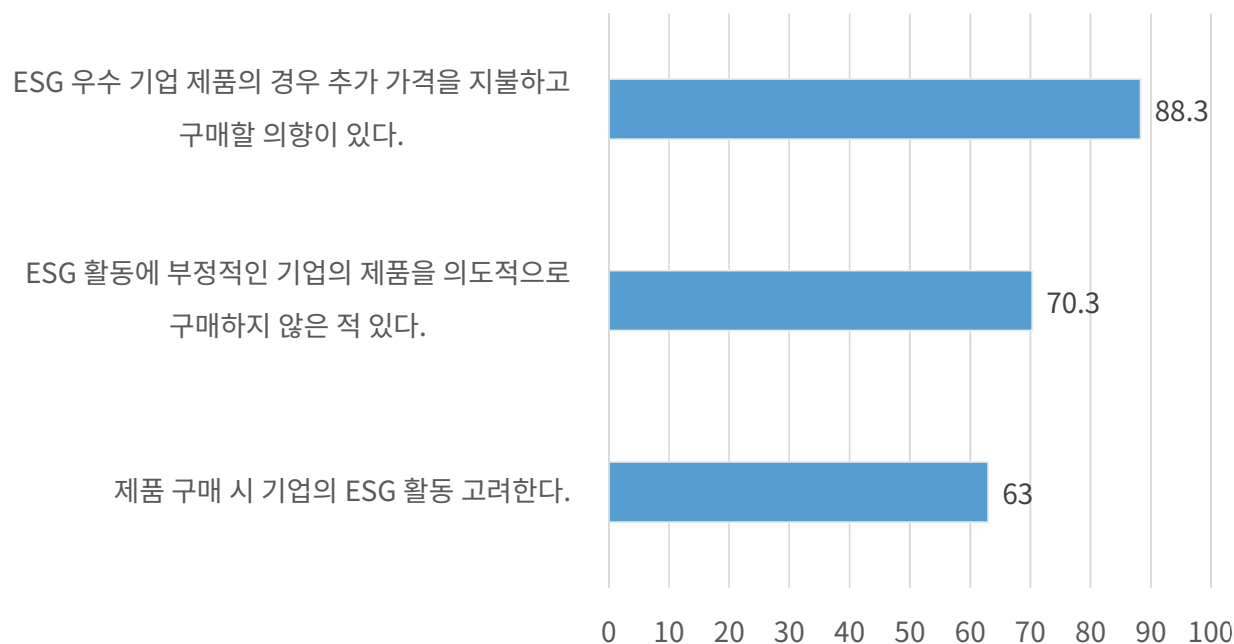
전통적 재무제표는 기업의 가치창출 역량 정보를 모두 반영하는 것은 아님

시장가치(P)와 장부가치(B) 간 차이의 상당 부분은 ESG 요소에 기인할 가능성 大

재무적 요인에 의존하던 기존 기업 분석 방식은 기업의 지속가능성 설명 부족

→ 기업의 비재무적 요인인 ESG 활동 반영 필요성 증대

ESG 경영에 대한 국민인식조사 결과 (대한상의, 21.5)



지속가능성 관련 정보(ESG정보)가 경영활동에 미치는 영향이 높아짐



개별 기업의 상황 및 전략에 맞는 ESG 경영이 요구됨

모든 기업들은 의무적인 방식의 ESG가 아닌, 기업의 경영전략 및 산업 여건에 맞는 ESG를 추구하는 방향으로 나아가야 함

“진정한 ESG 경영 시스템이 구축되기 위해서는 기업의 비전과 목표, 전략 등 경영 체계 전반에서 ESG 맞춤 재설계 작업이 필요”

(출처:한국거래소)

그러나, 기업들은 ESG 전략을 수립하는 데 난항을 겪고 있음

IFRS(국제회계기준위원회)는 2030년까지 상장사들의 ESG정보 공시를 의무화하여 투자자들의 더 나은 의사결정 지원 예정

금융위원회 또한 2025년부터 ESG의무공시를 도입할 예정이었으나, ‘준비기간이 필요하다’는 기업의 요청에 따라 1년 유예

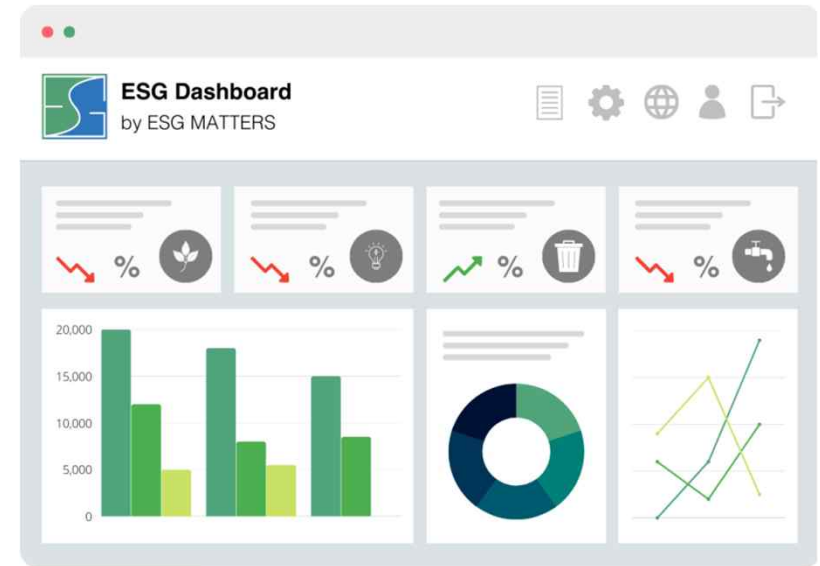
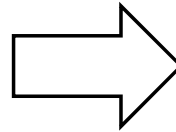
홈 > 증권 > 증권일반

[단독] "삼성도 공시 준비 안돼"...시스템 갖춘 곳 14%뿐

입력 2023-09-08 17:36:11 수정 2023.09.08 19:16:55 서종갑 기자

(출처: 서울경제)

통용되는 지표로서의 ESG가 아닌 개별 기업 맞춤형 ESG가 요구되나,
기업들은 맞춤형 전략 수립에 어려움을 겪고 있음



해당 기업의 공시정보에서 ESG가 실제로 얼마나 중요한 역할을 하는지를
분석하고 시각화하는 자동화 시스템을 제안하고자 함

02 프로젝트 정보

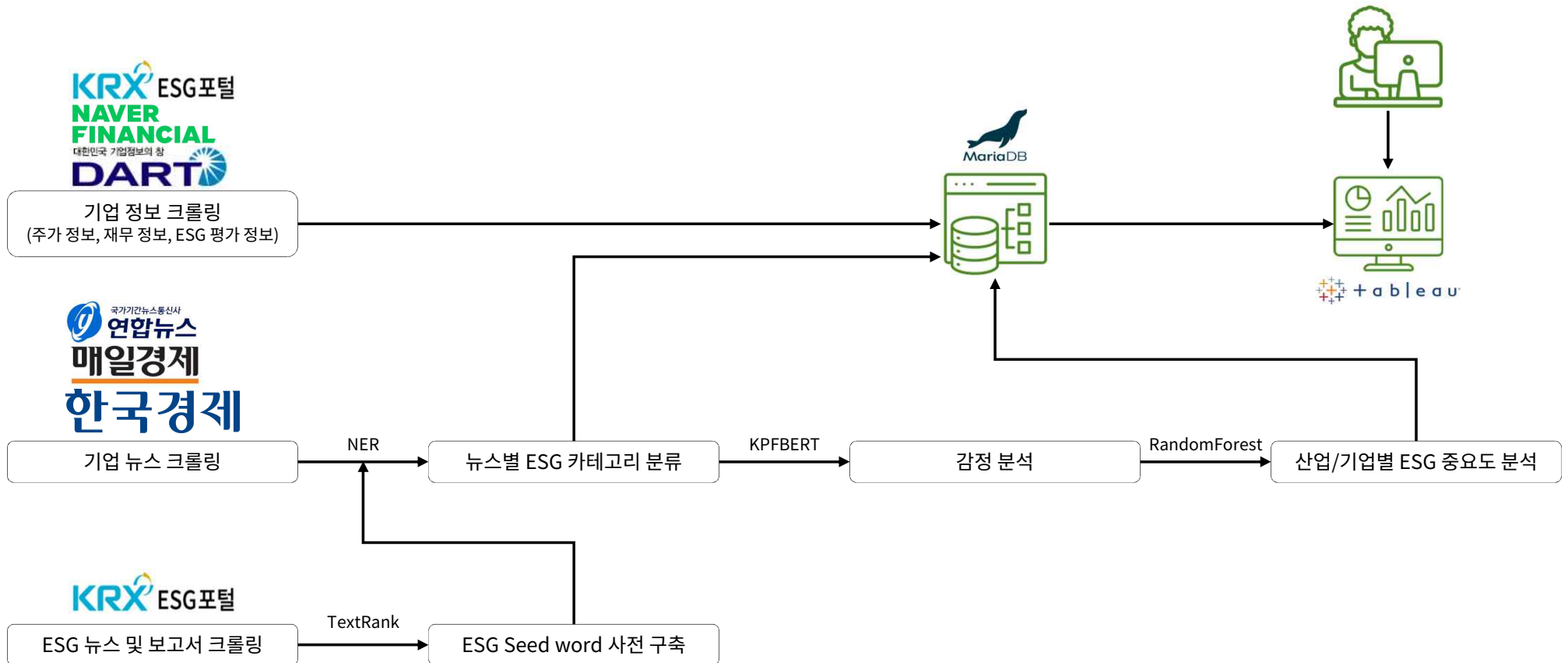
프로젝트 기간



업무 및 담당자	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
기업 정보 수집 (주가, 재무, ESG 평가) A, D						
기업 뉴스 수집 손준현, B						
ESG 뉴스 및 보고서 수집 손준현, D						
ESG Seed word 사전 구축 C, D						
ESG 카테고리 분류 B, C, D						
감정 분석 A, C						
산업/기업별 ESG 중요도 분석 A, D						
DB 구축 및 데이터 저장 손준현, B						
대시보드 디자인 손준현, A, B						

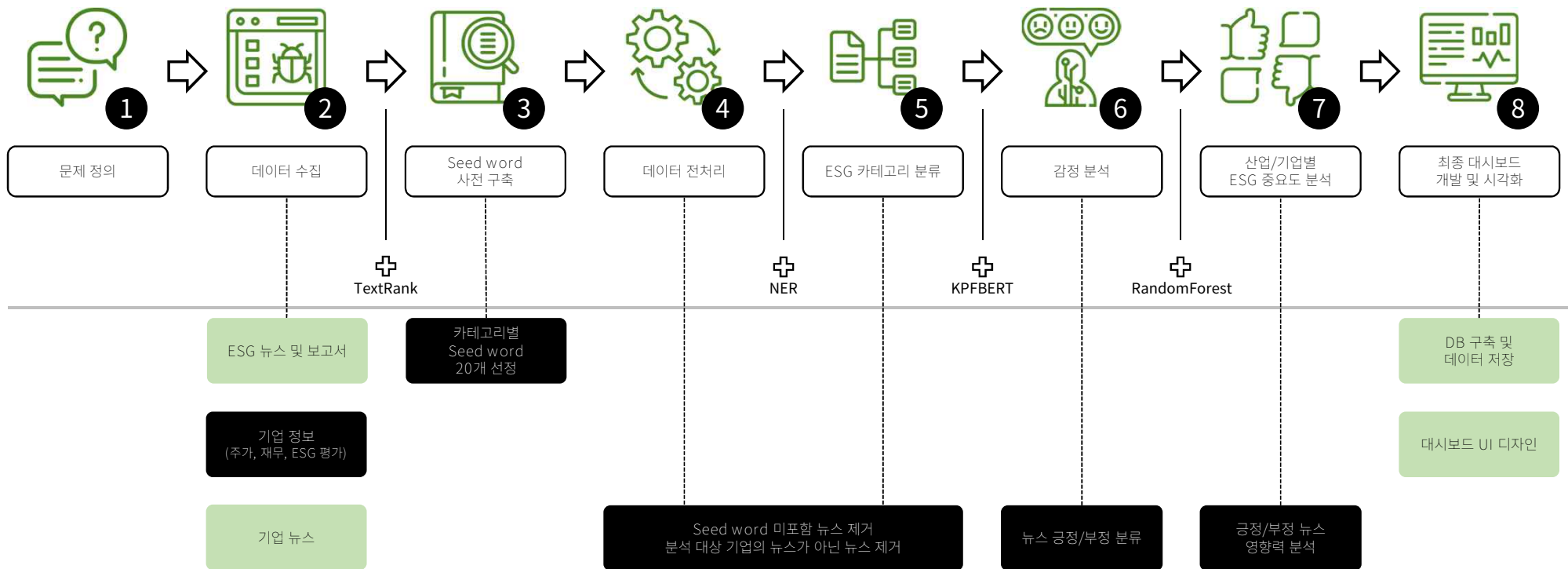
02 프로젝트 정보

서비스 플로우

















02 프로젝트 정보

프로젝트 진행 프로세스





Language	 Python					
Framework / Library	 Selenium	 Scrapy	 Pytorch	 Pandas	 Matplotlib	 Scikit-learn
IDE & Communication	 Visual Studio Code	 Google Colab	 Notion	 Slack	 Figma	
DBMS	 MariaDB					
BI	 Tableau					

02 프로젝트 정보

주요 기능 구현 상세 데이터 수집

기업 정보 (주가, 재무, ESG평가)

일별시세		
날짜	종가	전일비
2023.10.30	67,600	▲ 300
2023.10.27	67,300	▲ 600
2023.10.26	66,700	▼ 1,300
2023.10.25	68,000	▼ 500
2023.10.24	68,500	▲ 100
2023.10.23	68,400	▼ 400
2023.10.20	68,800	▼ 700
2023.10.19	69,500	▼ 1,000
2023.10.18	70,500	▲ 1,100
2023.10.17	69,400	▲ 2,100

기업실적분석		최근 연간 실적				최근 분기 실적				단위: 천원
주요재무정보		2020.12	2021.12	2022.12	2023.12(3)	2022.06	2022.09	2022.12	2023.03	2023.06
매출액(백만원)		2,368,070	2,796,048	3,022,314	2,696,291	772,036	767,817	704,646	637,454	609,055
영업이익(백만원)		359,939	531							
당기순이익(백만원)		264,078	39							
영업이익률(%)		15.20								
순이익률(%)		11.15								
ROE(%)		9.98								
부채비율(%)		37.07								
당차비율(%)		214.82								
유보율(%)		30,692.79	33.3							
EPS(원)		3,941								
PER(배)		21.09								
BPS(원)		39,406	4							
PBR(배)		2.06								
주당배당금(원)		2,995								
시가총액(백만원)		4,607								
배당성향(%)		77.85								

연결 재무상태표		
제 55 기 반기말 2023.06.30 현재		
제 54 기말 2022.12.31 현재		
(단위 : 백만원)		
자산	제 55 기 반기말	제 54 기말
유동자산	203,975,373	218,470,581
현금및현금성자산	79,919,775	49,680,710
단기금융상품	17,180,130	65,102,886
단기상각후평가금융자산		414,610
단기당기손익-공정가치금융자산	25,310	29,080
매출채권	36,104,142	35,721,563
미수금	5,822,772	6,149,209
선급비용	3,214,408	2,867,823
채고자산	55,504,798	52,187,866
기타유동자산	6,204,038	6,316,834
비유동자산	244,025,179	229,953,926
기타포괄손익-공정가치금융자산	10,308,125	11,397,012
당기손익-공정가치금융자산	1,573,457	1,405,468
관계기업 및 공동기업 투자	11,334,899	10,893,869
유형자산	177,869,942	168,045,398
무형자산	23,430,234	20,217,754
순확정급여자산	5,143,048	5,851,972
이연법인세자산	5,527,377	5,101,318
기타비유동자산	8,898,297	7,041,145
자산총계	448,000,552	448,424,507

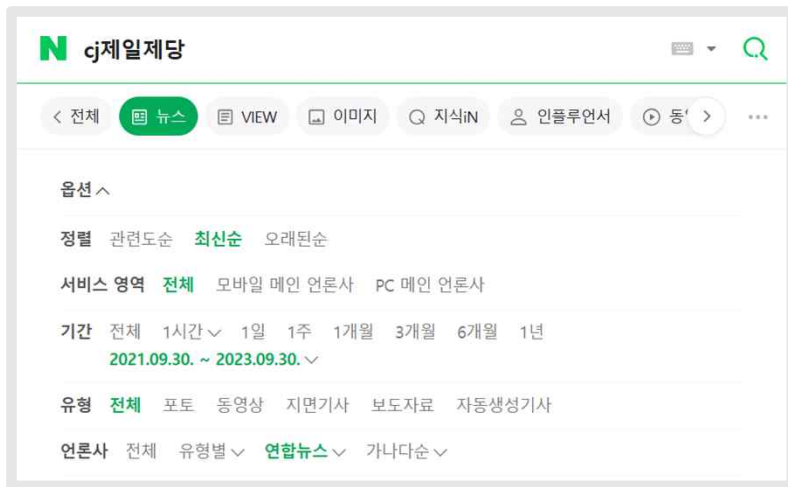
기업지배구조 공시우수법인										현재 791 사
회사명	KCGS				한국ESG 연구소	서스틴 베스트	Moody's	MSCI	S&P	지속가능경영 보고서공시여부
	종합등급	환경	사회	지배구조						
AJ네트웍스	B+	C	B+	B+	-	-	-	-	-	YES
AK홀딩스	B+	B	A	B	-	-	-	-	6	YES
BGF	B+	A	A+	B	-	-	-	-	18	YES
BGF리테일	A	A	A+	A	A	AA	23	BBB	58	YES
BNK금융지주	A	A	A	A	A	A	37	-	43	YES
BYC	D	D	D	D	-	-	-	-	-	YES
CJ	B+	B+	A	B+	A	AA	47	A	29	YES
CJ CGV	B+	B	B+	B+	-	-	-	-	10	YES
CJ대한통운	B+	B+	A+	B	B+	BB	39	-	62	YES
CJ푸드	B	C	B	B	-	-	-	-	-	YES

- 대상 기업 및 산업: 코스피 시가 총액 상위 100개 기업(지주사 제외)과 해당 기업들이 속한 산업 (2023. 09. 30. 기준)
 - ESG 의무공시를 우선적으로 시행하고, 원활한 뉴스 수집을 위해 언론에 충분히 노출이 되는 기업을 대상 기업으로 선정
- 수집 기간: 2021. 09. 30. ~ 2023. 09. 30.

02 프로젝트 정보

주요 기능 구현 상세 데이터 수집

기업 뉴스



- 수집 기간: 2021. 09. 30 ~ 2023. 09. 30
- 수집 기간 내 **단순 상호명 변경** 또는 **M&A로 인한 상호명 변경**과 같은 경우들을 조사하여 수집 범위에 포함시키는 작업 수행
- 수집 건수: 107,560건

ESG 뉴스 및 보고서



KRX에서 ESG로 분류된 뉴스 13,864건



2022년 발간된 지속경영보고서 110개

02

프로젝트 정보

주요 기능 구현 상세

Seed word 사전 구축

ESG Seed word 사전 구축

목적

각 기업의 ESG 보고서 및 ESG 뉴스의 내용을 추출하여
카테고리별 핵심 단어를 선정하고자 함

사용한 오픈 소스 및 활용

TextRank 패키지 ([GitHub](#))

단어 그래프를 생성하고,
생성된 그래프에 PageRank를 학습시켜 키워드를 추출

Why TextRank? - 타 알고리즘과의 비교

KeyBERT - 타 문서와 비교해 특이한 단어들을 추출하는 알고리즘 특성 상 ESG 정보가 담긴 문서의 중요 단어들을 추출하기 어려움

KR-WordRank - 한국어에 특화된 WordRank 알고리즘이나, '기후변화', '취약계층' 등의 복합어 추출 불가

02 프로젝트 정보

주요 기능 구현 상세 Seed word 최종 선정 결과

환경(Environment)

키워드	점수
친환경	19.64
물질	16.14
환경경영	13.75
폐기물	13.18
에너지	12.01
배출	9.43
화학	8.10
인증	7.61
저감	7.33
배출량	7.07
기후변화	6.84
사용량	6.59
감축	6.06
적용	5.96
온실가스	5.82
자원	5.51
구매	5.26
용수	5.26
처리	5.03
탄소중립	4.85

사회(Social)

키워드	점수
노조	34.09
화재	22.59
파업	20.22
노동자	17.18
작업	14.61
오후	14.08
소송	13.43
대표	12.13
데이터센터	11.31
이사회	10.83
조합원	10.80
근로자	10.37
검찰	10.35
원인	10.18
임금	10.10
프로그램	9.35
택시	9.24
교섭	9.23
예정	8.98
이용	8.86

지배구조(Governance)

키워드	점수
의혹	27.09
구속	23.59
횡령	20.42
부지사	16.42
영장	14.84
기소	13.28
검사	13.10
변호사	12.92
지검	11.44
부장	11.07
해외	10.95
범죄	10.49
재판	10.09
대북	9.97
송금	9.16
자금	9.09
주주	8.76
뇌물	8.65
배임	8.64
특검	8.42

02

프로젝트 정보

주요 기능 구현 상세

데이터 전처리

Seed word가 포함되지 않은 기사 제거

```
● ● ●  
  
# seed word 목록  
seed_word = '''친환경  
폐기물  
...  
특검'''  
  
seed_word = '|'.join(seed_word.split('\n'))  
  
# seed word가 포함되지 않은 뉴스 제거  
news = news[news['contents'].str.contains(seed_word)==True]
```


02

프로젝트 정보

주요 기능 구현 상세

ESG 카테고리 분류

개체명 인식(NER) 적용

목적

문장에서 개체명(인명, 기관명), 수량 표현(통화, 비율), 날짜 표현(날짜, 시간, 기간) 등의 개체를 인식하여 일상적 표현과 구분하고자 함

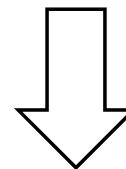
사용한 오픈 소스 및 활용

BERT 기반의 한국어 개체명 인식기 ([GitHub](#))

Pre-trained 데이터 내 B-ORG 위치에 분석 대상 기업명을 삽입 후
총 3번의 학습을 진행함

<NER 적용 예시>

“선진(총괄사장 이범권)이 비육우 사료 신제품 ‘포텐셜9’과 ‘에센셜K’를 출시, 한우 경쟁력 강화에 기여 하겠다고 밝혔다.”



[('선진', 'B-ORG'), ('(', 'O'), ('총괄', 'O'), ('사장', 'O'), ('이범권', 'B-PER'), ('(', 'O'), ('이', 'O'), ('비', 'O'), ('육우', 'O'), ('사료', 'O'), ('신제품', 'O'), ('(', 'O'), ('포텐셜', 'B-POH'), ('9', 'I-POH'), ('(', 'O'), ('과', 'O'), ('(', 'O'), ('에센셜', 'B-POH'), ('K', 'I-POH'), ('(', 'O'), ('를', 'O'), ...

02 프로젝트 정보

주요 기능 구현 상세 최종 데이터셋

감정 분석에 사용될 최종 뉴스 데이터

	press	date	company	title	contents
0	연합뉴스	2023.09.20. 오전 10:52	SK오션플랜트	HLB, 코스닥→코스피 이전 추진..."상장주선인 계약 체결"	(서울=연합뉴스) 송은경 기자 = 코스닥 상장사 HLB는 20일 유가증권시장으로 이...
1	연합뉴스	2023.08.28. 오후 4:53	SK오션플랜트	엘앤에프, 유가증권시장 이전 상장 추진...임시주총서 확정	엘앤에프 CI[엘앤에프 제공] (서울=연합뉴스) 송은경 기자 = 코스닥 상장사...
2	연합뉴스	2023.09.11. 오후 5:00	SK오션플랜트	'취업 역량 강화' 경남 주요 대학, 청년에게 다양한 정보 제공	창원대·경남대·경상국립대·인제대, 기업 연결 채용정보 마련 \n\n\n채용정보 ...
3	연합뉴스	2023.09.11. 오전 10:34	SK오션플랜트	SK에코플랜트, 환경·에너지 분야 혁신기술 발굴 공모전	'2023 콘테크 미트업 데이'[SK에코플랜트 제공. 재판매 및 DB 금지](서울=...
4	연합뉴스	2023.08.31. 오전 9:51	SK오션플랜트	SK에코플랜트, 스타트업 기술공모전...소프트베리 등 3개사 수상	스타트업 혁신기술 발굴해 사업화 지원\n\n\n'테크 오픈 컬래버레이션' 시상식...
...
36637	한국경제	2022.02.09. 오전 11:02	유한양행	"도급·협력업체 위험성 평가도 필수...개선조치 기록 5년 보존해야"	"도급·협력업체 사업주도 수급업체와 함께 위험성 평가를 반드시 해야합니다." 신인재...
36638	한국경제	2022.01.25. 오후 3:23	유한양행	유한양행, 직원 기부로 소아암 치료비 지원...어린이 과학교실·청소년 진로교육	유한양행은 비대면으로 온라인 임직원 경매를 진행했다. 수익금은 한국백혈병어린이재단에...
36639	한국경제	2022.01.11. 오후 3:10	유한양행	화이자 '먹는 코로나 치료제' 14일부터 처방...유한양행 유통	13일부터 약국 및 치료센터 공급 예정정부, 내일 활용방안 등 발표화이자의 먹는(경...
36640	한국경제	2021.12.06. 오후 12:02	유한양행	[하이서울기업을 보다] "엄마, 왜 주방세제로 샴워해?"...천연에 진심인 아세즈	박성희 화인코리아 코퍼레이션 대표 인터뷰친환경 제품 통해 깨끗한 자연 만들기 '앞장...
36641	한국경제	2021.11.02. 오후 5:19	유한양행	에이비엘바이오, 이중항체 기반 면역항암제 임상 확대	'ABL103' 등 파이프라인 늘려면역항암제 임상 2년내 3→6 건2018년 미국 트...

02 프로젝트 정보

주요 기능 구현 상세 감정 분석

뉴스 감정 분석

사용한 오픈 소스

언론진흥재단 BERT(KPFBERT) ([GitHub](#))

한국형 표준 뉴스 기사 인공지능 언어 모델로,

약 4,000만 건의 기사를 학습해 언론사 및 뉴스 기사 활용에 최적화

기업명과 Seed word를 포함한 문장 기준으로 뉴스 분리

<감정 분석 예시>

"현대자동차가 전기차 고객 대상으로 구독형 충전 요금제 '럭시패스 H'를 선보인다고 16일 밝혔다. 고객의 전기차 충전 요금 부담을 줄이고 전기차 접근성을 높이기 위해서다. 럭시패스 H에 가입한 고객은 매달 일정액을 내면 약정한 충전량 한도 내에서 충전 요금을 할인받을 수 있다. 앞서 테슬라가 구매 고객에게 충전소인 슈퍼차저 1년 무료 정책을 펴다가 폐지한 바 있다. ..."

["현대자동차가 전기차 고객 대상으로 구독형 충전 요금제 '럭시패스 H'를 선보인다고 16일 밝혔다.[SEP]고객의 전기차 충전 요금 부담을 줄이고 전기차 접근성을 높이기 위해서다.", '고객의 전기차 충전 요금 부담을 줄이고 전기차 접근성을 높이기 위해서다.[SEP]럭시패스 H에 가입한 고객은 매달 일정액을 내면 약정한 충전량 한도 내에서 충전 요금을 할인받을 수 있다.', '럭시패스 H에 가입한 고객은 매달 일정액을 내면 약정한 충전량 한도 내에서 충전 요금을 할인받을 수 있다.[SEP]앞서 테슬라가 구매 고객에게 충전소인 슈퍼차저 1년 무료 정책을 펴다가 폐지한 바 있다.', ...]

문장들을 2문장씩 묶어서 감정 분석 진행
(0: 부정, 1: 중립, 2: 긍정)

(tensor([[[6.9349e-05, 9.6970e-05, 9.9983e-01]]], '긍정'))

02 프로젝트 정보

주요 기능 구현 상세 산업/기업별 ESG 중요도 분석

RandomForest

특징

앙상블 학습(Ensemble Learning)의 일종으로,
다수의 의사 결정 트리(Decision Tree)를 결합하여 예측 모델을 생성

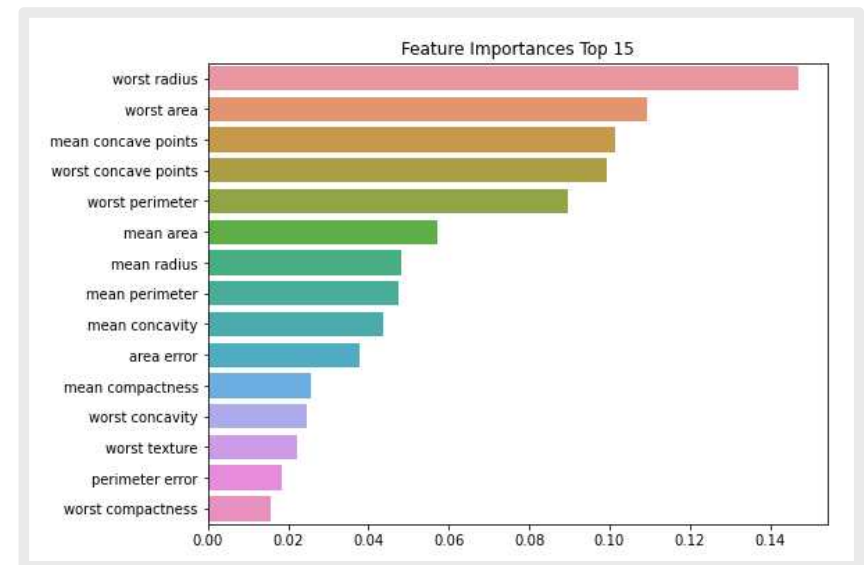
변수의 중요도를 평가할 수 있어 특성 선택에 유용하게 사용됨

R2 Score를 통해 변수들의 설명력을 측정 가능

잔차 제곱합(SSE): 회귀 모델이 예측한 값과 실제 종속 변수 값 사이의 차이를 제곱하여 합한 값

총 제곱합(SST): 실제 종속 변수 값과 평균값 사이의 차이를 제곱하여 합한 값

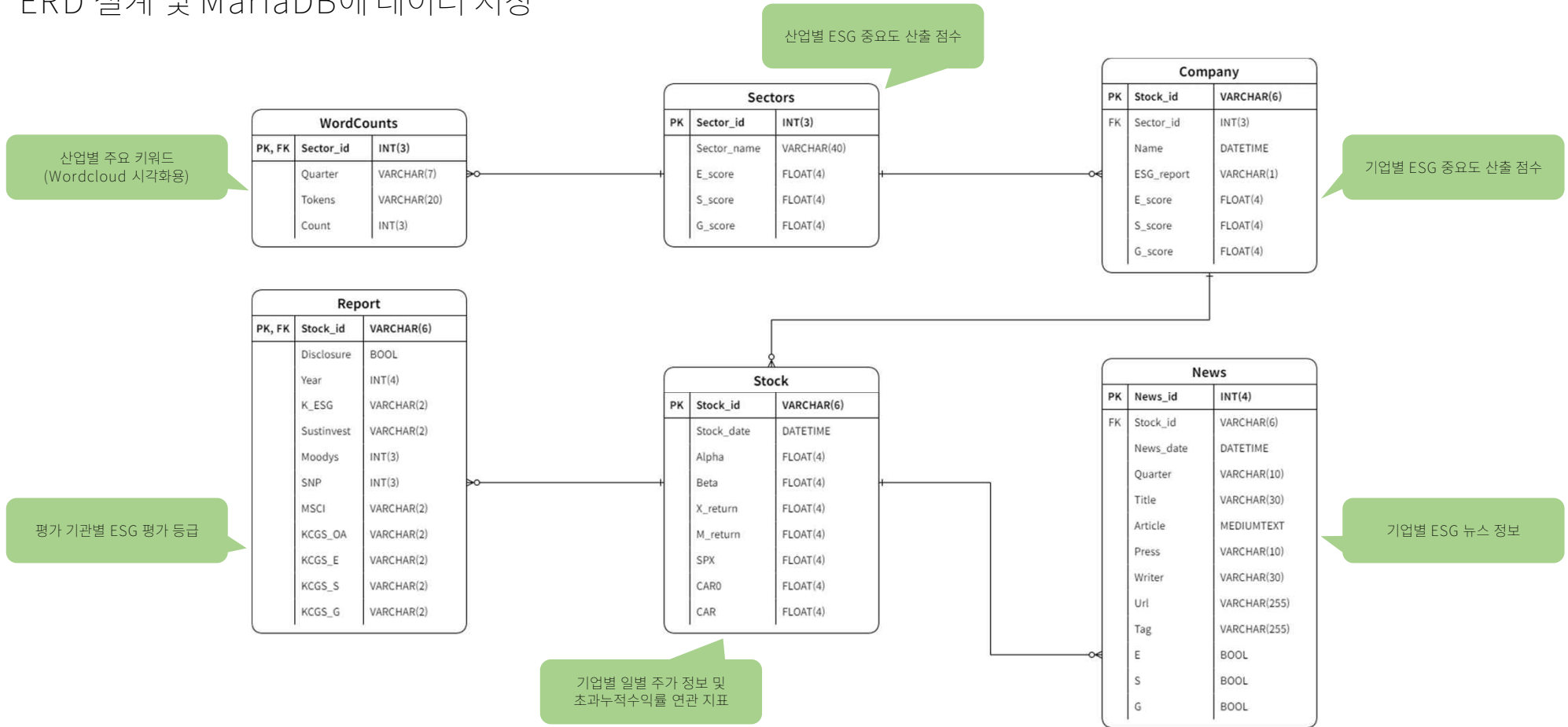
$$R^2 = 1 - (SSE / SST)$$



02 프로젝트 정보

주요 기능 구현 상세 데이터베이스 구축

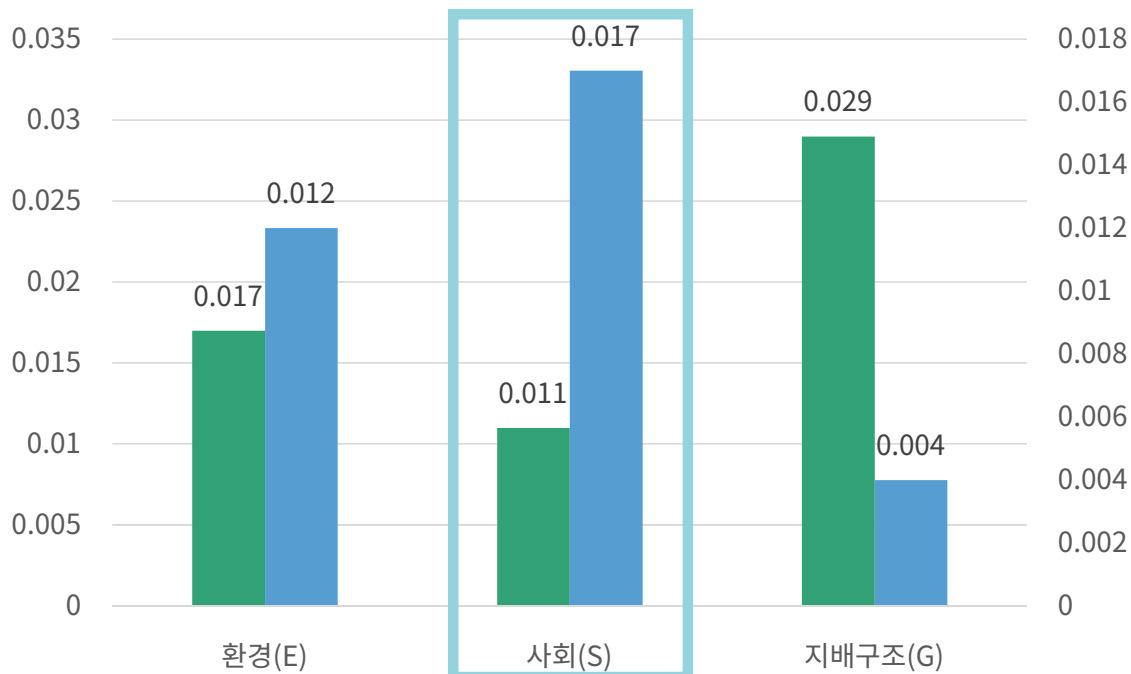
ERD 설계 및 MariaDB에 데이터 저장



Case 1 - 삼성전자

전자부품 컴퓨터 영상 음향 및 통신 장비 제조업 산업 및

삼성전자의 E,S,G 분야별 중요도



2023년 삼성전자 ESG평가

(평가기관: 한국ESG기준원)

환경(E)	A
사회(S)	A+
지배구조(G)	B+

분석 결과

삼성전자의 경우 사회(Social) 분야의 중요도가
산업 평균에 비해 높게 나타났음.

한국ESG기준원의 평가에 따르면,
삼성전자는 사회(Social)부분 ESG점수가 높으므로,
당사의 중요 영역을 파악하고 ESG를 잘 실천하고 있는 것으로 보임.

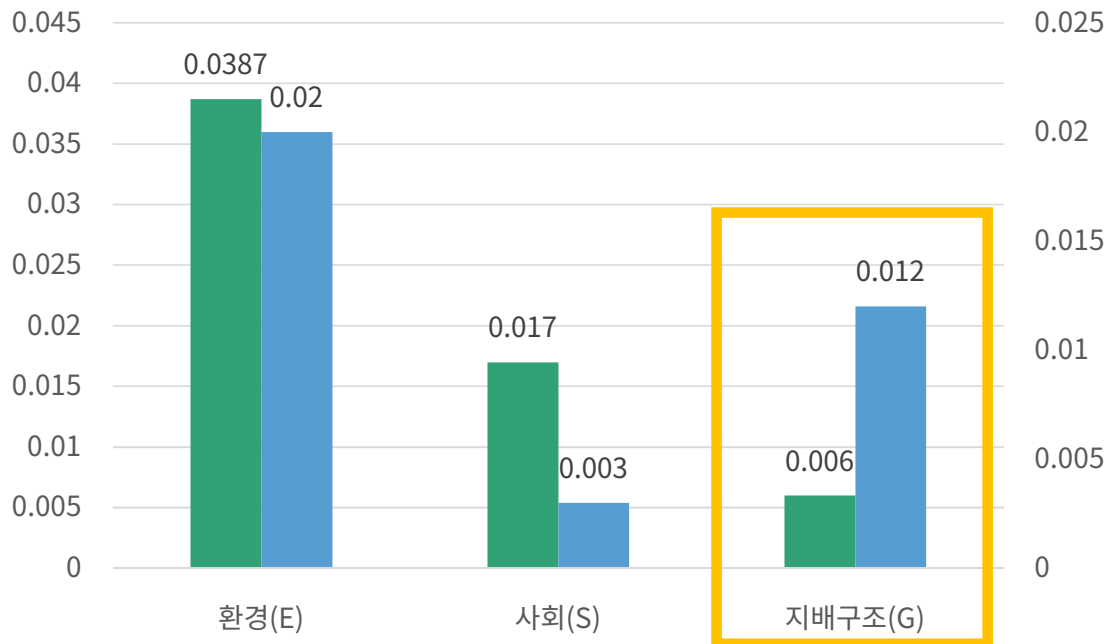
02

프로젝트 정보

성과 및 결과

Case 2 - 효성티앤씨

화학물질 및 화학제품 제조업 산업 및
효성티앤씨의 E,S,G 분야별 중요도



2023년 효성티앤씨 ESG평가

(평가기관: 한국ESG기준원)

환경(E)	A
사회(S)	A+
지배구조(G)	B+

분석 결과

효성티앤씨의 경우 지배구조(Governance) 분야의 중요도가 산업 평균에 비해 높게 나타났음.

하지만 한국ESG기준원의 평가에 따르면, 효성티앤씨는 지배구조(Governance)부분 ESG점수가 낮으므로, 지배구조 측면을 보완해야 함을 인지하고 ESG를 전략적으로 실천할 필요가 있어 보임.

02 프로젝트 정보

성과 및 결과

최종 대시보드 구축 결과

예시) 삼성전자



기업과 산업별 ESG 분야별 중요도 평가 과정에서 종속 변수 설정의 어려움

Problem

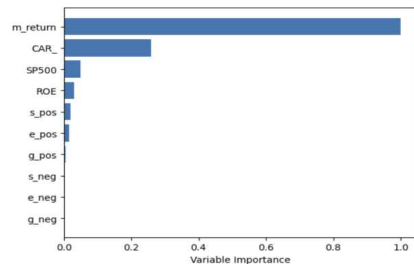
ESG 긍정/부정의 중요도와 연관된
기업 가치와 관련된 종속 변수 설정의 어려움



Solution

일별 주가 수익률의 변동을 종속 변수로 삼아
ESG 긍정/부정 뉴스의 영향력 분석

삼성전자 분석 결과

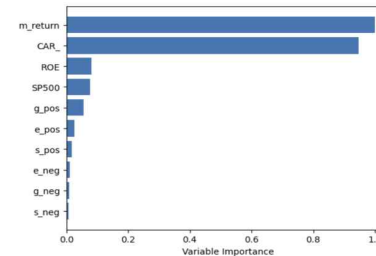


분석 요약

R2_score: 0.648

- | | |
|--------|--------|
| ① S_긍정 | ④ S_부정 |
| ② E_긍정 | ⑤ E_부정 |
| ③ G_긍정 | ⑥ G_부정 |

전자부품, 컴퓨터, 영상, 음향 및 통신 장비 제조업 분석 결과



분석 요약

R2_score: 0.372

- | | |
|--------|--------|
| ① G_긍정 | ④ E_부정 |
| ② E_긍정 | ⑤ G_부정 |
| ③ S_긍정 | ⑥ S_부정 |



환경 영향력 증가: 기업과 자본시장의 기후변화 리스크 대응

기후변화(E) 리스크가 자본시장 전체에 영향을 미침에 따라,
시장 전체 위험(Total Risk)에 관심이 많은 기관투자자들의 주요 관심사가 되었기 때문으로 짐작된다.

ESG 관련 보도의 주가 영향 분석과 재무적 요인의 복잡성

선행 연구 결과와 일치하게, ESG 관련 긍정적 보도가 주가에 미치는 영향을 확인할 수 있었다.
그러나 기업의 ESG 활동이 주가 수익률 외에도 재무제표 등의 재무적 요인에 미치는 영향을 파악하는 것은 복잡한 문제였다.
앞으로는 이러한 재무적 요인에 대한 더 깊은 이해를 바탕으로 분석을 보완해야 할 것이다.

기업별 뉴스 발행량 고려한 분석 방안 강구 필요성

분석 대상 기업을 다양하게 선택했으나, 기업 간 뉴스 발행량의 차이를 고려하지 않아서, 뉴스 발행량이 적은 기업의 경우 결과가 편향되었다.
향후 분석에서는 기업별 뉴스 발행량 차이를 고려하여 더욱 신뢰할 만한 결과를 얻도록 해야 할 것이다.

Tableau vs. Redash

		
장점	다양한 시각화 및 분석 기능 제공 다양한 데이터 소스와의 연동 지원 뛰어난 디자인 및 커스터마이징 기능	오픈 소스로 무료 이용 가능 유연한 쿼리 작성 및 실행 기능 웹 기반으로 접근이 용이함
단점	상대적으로 높은 비용 대규모 데이터 처리 시 성능 저하 서버 구축 및 유지보수에 별도 비용 필요	사용자 친화적인 UI 부재 제한된 시각화 및 분석 기능