1차 보고(01/09)

일정

1단계 (4일)	 텍스트 형식의 뉴스 본문을 벡터로 변환시킨다. 뉴스 간 유사도를 측정한다.
2단계 (2일)	 유사도를 기준으로 뉴스 본문 벡터를 2차원 벡터로 다시 변환한다. 2차원 벡터 기준으로 유사한 뉴스 그룹을 모은다.
3단계 (3일)	각 그룹별 중심이 되는 뉴스를 찾는다.그룹 안의 중심이 되는 뉴스와 유사하지 않은 뉴스를 찾는다.
4단계 (3일)	 유사하지 않은 뉴스 제거 전과 후를 비교한다. 시각화를 위해 2차원 상 데이터 분포 시각화하여 비교한다.
4단계 (1일)	- 최종 결과물 분석 및 중간보고서 작성
5단계	- 1월 31일 중간 결과 보고

2차 보고(01/16)

2023년 1월 9일 미팅 내용

- 계획서 내용 대로 2일 안에 결과 내기. -> 다음주로 연기,
- 2일안에 가능하도록, 샘플수를 줄여서 실험 결과 가지고 오기. -> 기간 연장으로 추가 실험 진행

데이터

뉴스 본문 사용

기간: 2018-01-01 ~ 2021-04-28

카테고리: 대기오염, 환경관리체계로 실험. (48시간 내 가능한 실험 구성을 위해 세팅)

뉴스 수

- 대기오염: 25개
- 환경관리체계:25개

총 단어 수 :

- 대기오염: 4286개
- 환경관리체계: 4550개

차별성

- 2018~2022년 기준 약 87600개의 논문에서 이와 같은 방법을 적용한 사례 없음.

출처 : arXiv 통계

뉴스 본문 벡터 표현

1단계 : 단어의 pool과 같은 길이의 행렬을 만든다.

문서1: 먹고 싶은 사과

문서2: 먹고 싶은 바나나

문서3: 길고 노란 바나나 바나나

문서4: 저는 과일이 좋아요

	7	과일이	길고	노란	먹고	바나나	사과	싶은	저는	좋아요
	문서1	0	0	0	1	0	1	1	0	0
3	문서2	0	0	0	1	1	0	1	0	0
	문서3	0	1	1	0	2	0	0	0	0
	문서4	1	0	0	0	0	0	0	1	1

TF-IDF 계산법

2단계: 행렬의 값은 각 단어의 TF-IDF 값을 넣는다.

[TF-IDF 구하는 방법]

한 뉴스에서 나온 단어 빈도 / 전체 뉴스에서 나온 단어 빈도

(예시) 단어 = 환경

한 뉴스에서 나온 "환경" 수 / 전체 뉴스에서 나온 "환경" 수

TF-IDF로 표현

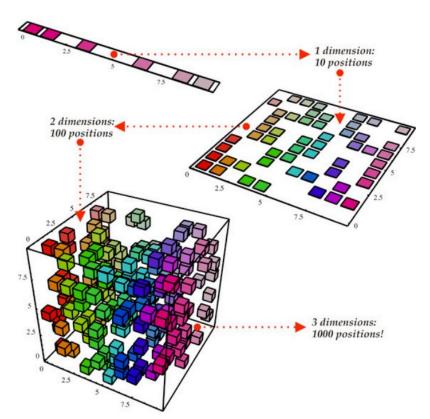
2단계: 행렬의 값은 각 단어의 TF-IDF 값을 넣는다.

	과일이	길고	노란	먹고	바나나	사과	싶은	저는	좋아요
0	0.00000	0.00000	0.00000	0.526405	0.00000	0.667679	0.526405	0.00000	0.00000
1	0.00000	0.00000	0.00000	0.577350	0.57735	0.000000	0.577350	0.00000	0.00000
2	0.00000	0.47212	0.47212	0.000000	0.74445	0.000000	0.000000	0.00000	0.00000
3	0.57735	0.00000	0.00000	0.000000	0.00000	0.000000	0.000000	0.57735	0.57735



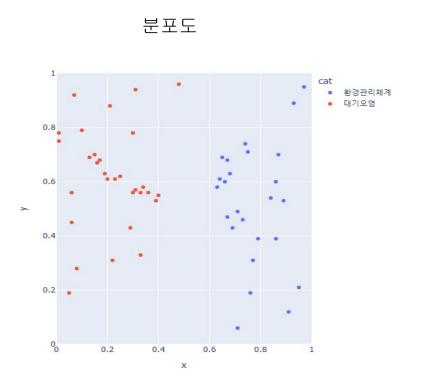
둘의 차이를 줄이는 방식으로 훈련

고차원 벡터를 2차원 벡터로

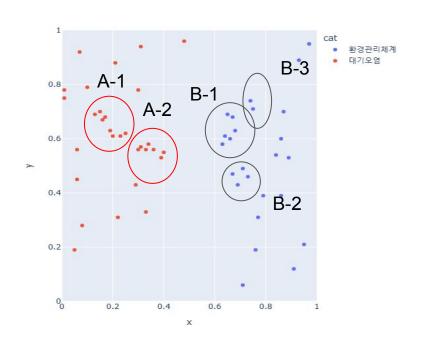


상자 = 뉴스 한개

본문-> 가장 작은 의미 단위로 분해 -> 벡터화 -> 2차원 공간에 시각화



분포도: 그룹 분석



그룹별 중심단어 표시

그룹 A -1 [대기오염]

공통 단어 : 오염, 배출, 대기, 물질

뉴스 예시:

- 1. '대기오염 측정치 조작' 영풍제련소 임원, 징역 1년 2개월 실형 | 뉴스민 (newsmin.co.kr)
- 2. 현대제철, 방지시설 없이 오염물질 배출 매일경제 (mk.co.kr)
- 3. 대한유화 온산공장 강력 처벌,환경오염에 경고 메시지 <종합 < 사회 < 기사본문 경상일보 (ksilbo.co.kr)
- 4. <u>롯데케미칼, 대기오염물질 초과배출 부과금 취소 소송 "패소" < 경제 < 기사본문 데일리시사닷컴</u> (dailysisa.com)
- 5. "여수산단 배출조작 기업들 사회공헌 이행의지 있나" 노컷뉴스 (nocutnews.co.kr)
- 6. '무용제 도료' 개발로 선박 도료업계 '친환경' 활기 아시아경제 (asiae.co.kr)
- 7. 정부, 화학물질 배출량 산정방법 개선 착수...'배출조작' 잠재울까 이투데이 (etoday.co.kr)
- 8. 전남도의회 "여수산단 환경개선 기업 책임 다해야" 노컷뉴스 (nocutnews.co.kr)
- 9. <u>충남도, 대기오염물질 감축 5개 기엄 친환경사업장 선정 | 연함뉴스 (yna.co.kr)</u>

그룹 A -2 [대기오염]

공통 단어: 미세먼지, 산화물,질소 U 석탄, 가스

뉴스 예시:

- 1. 포스코, 친환경 청정설비로 미세먼지 감축…질소산화물 최대 80% 절감 http://www.economytalk.kr/news/articleView.html?idxno=209194
- 2. 한국가스공사, LNG공급 합작회사 매출 10년내 1조 달성

 http://mbiz.heraldcorp.com/view.php?ud=20201124000172&ntn=0
- 3. 세아STX엔테크, 보령화력발전소 대기환경개선 설비 본격 가동해 https://www.mk.co.kr/news/business/view/2020/11/1230410/
- 4. 그린 일자리 최우수상 포스코에너지 미세먼지 종합대책 선제 대응...채용 확대 앞장 http://mbiz.heraldcorp.com/amp/view.php?ud=20201203000425
- 5. 한국지역난방공사, 대구 열병합발전소 친환경에너지 개선사업 주민설명회 개최 http://www.newsci.com/news/articleView.html?idxno=805308
- 6. <u>원전·석탄화력 대체한다는 LNG발전소... 전국서 잇따라 '좌초' (naver.com)</u>
- 7. 현대제철, 친환경 제철소로 탈바꿈 (g-enews.com)

그룹 B -1 [환경관리체계]

공통 단어: 뉴딜,친환경,녹색,책임,채권,영향,강화,긍정,가치,미래,감사,성과,확대,역량

뉴스 예시 :

文정부 '그린 뉴딜' 추진...은행권 친환경 사업 탄력 받나

http://m.kukinews.com/m/m_article.html?no=789382

녹색채권 발행, 재생에너지 투자... '환경'에 속도 내는 금융업계

https://news.chosun.com/site/data/html_dir/2020/08/17/2020081701479.html

민주, '그린뉴딜' 힘 보탤 녹색금융공사 설립 추진

https://www.news1.kr/articles/?4084357

[2020 국감] 금융사 책임투자·녹색금융 규모 51조원...분류체계 확립해 옥석가려야

- 매일경제

https://www.mk.co.kr/news/economy/view/2020/10/1040243/

NH농협금융, 환경부와 금융사 최초 MOU 체결

http://www.wiseenews.com/news/articleView.html?idxno=32925

[ESG경영] NH·삼성·KB·미래, 뉴딜 맞춰 ESG채권 발행 앞다퉈

https://www.fntimes.com/html/view.php?ud=20210227113047209dd55077bc2_18

그룹 B -2 [환경관리체계]

공통 단어: 과징금, 관리, 희망퇴직, 희생, 규제, 강화, 환경

뉴스 예시 :

[2020 국감 이슈] 화력발전 5개사, 5년간 환경오염 관련 위반 94 ...

http://www.newslock.co.kr/news/articleView.html?idxno=31817

[2020국감] 화력발전 5사, 환경규제 위반 심각...최근 5년간 위반 94건

http://www.straightnews.co.kr/news/articleView.html?idxno=84913

화력발전 5사. 환경법령 위반 심각

http://www.hkbs.co.kr/news/articleView.html?idxno=594464

이 와중에... 르노삼성·쌍용차 300억대 과징금까지 - 조선일보

https://www.chosun.com/economy/auto/2021/03/04/K6XVBYZ5OFHSTANRYSL3U4Z2MY/

그룹 B -3 [환경관리체계]

공통 단어: 농업,유통,온실가스,녹색,동물,친환경,환경친화,실천,상생,지역

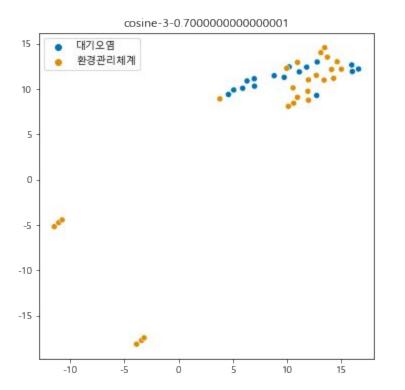
뉴스 예시:

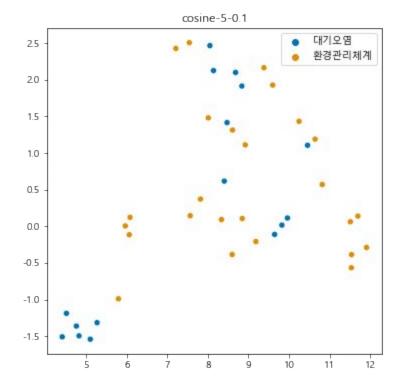
의식주책(衣食住冊)에 부는 '친환경' 바람

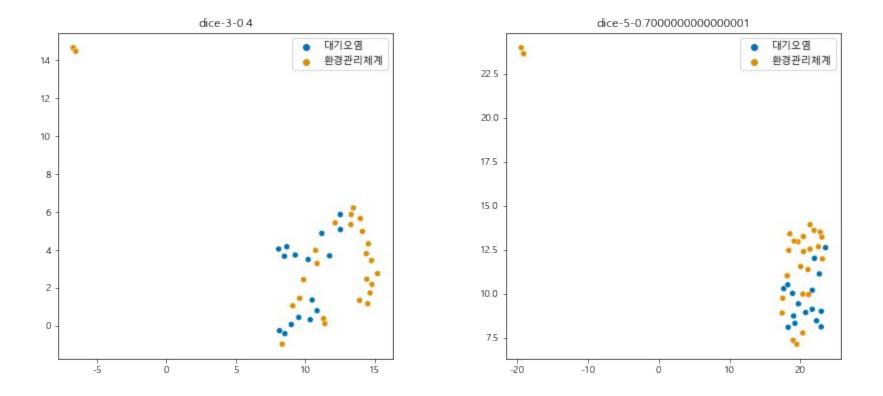
https://www.readersnews.com/news/articleView.html?idxno=100349

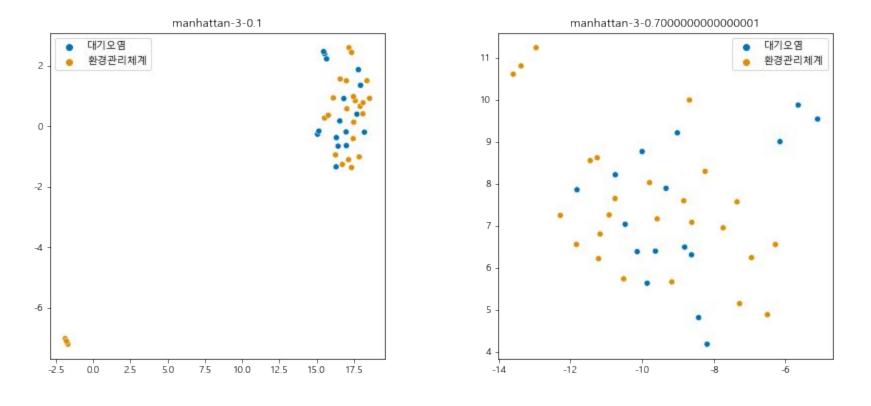
풀무원 올가홀푸드, 친환경 농업 발전 공로 '농림축산식품부 장관 ...

https://www.koreapost.co.kr/news/articleView.html?idxno=55570









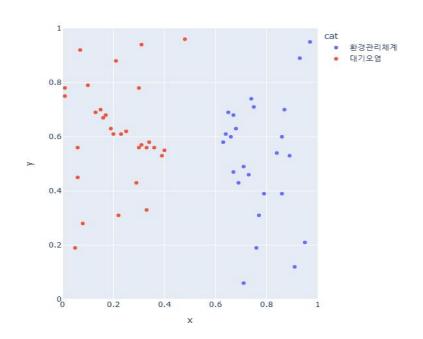
3차 보고(2/1)

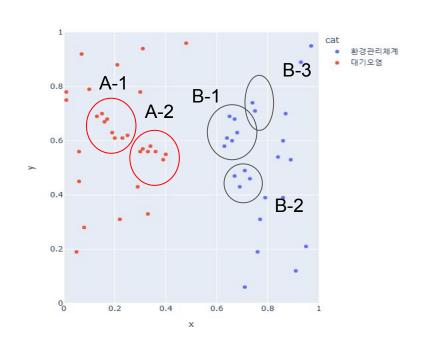
2023년 1월 16일 미팅 내용

- 2차때 사용한 방법으로 계속 진행하라.

2차 발표 recap 본문-> 가장 작은 의미 단위로 분해 -> 벡터화 -> 2차원 공간에 시각화

분포도 : 그룹 분석





그룹별 중심단어 표시

데이터

뉴스 본문 사용

기간: 2018-01-01 ~ 2021-04-28

뉴스 수

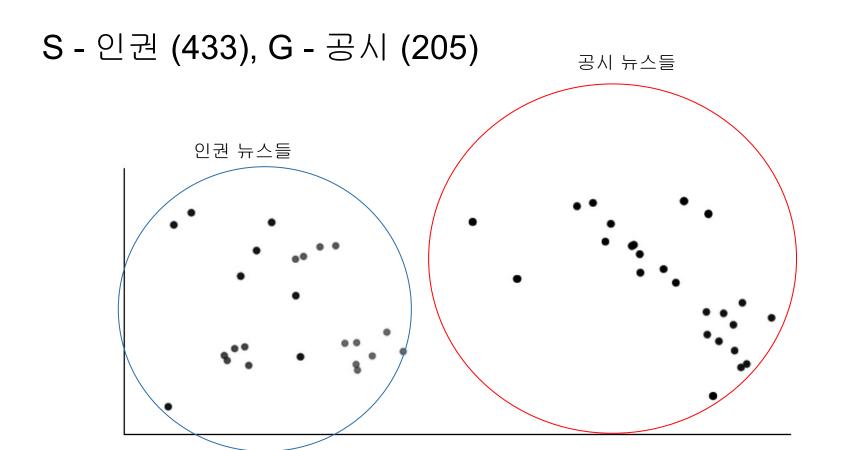
- 공시:25개 G

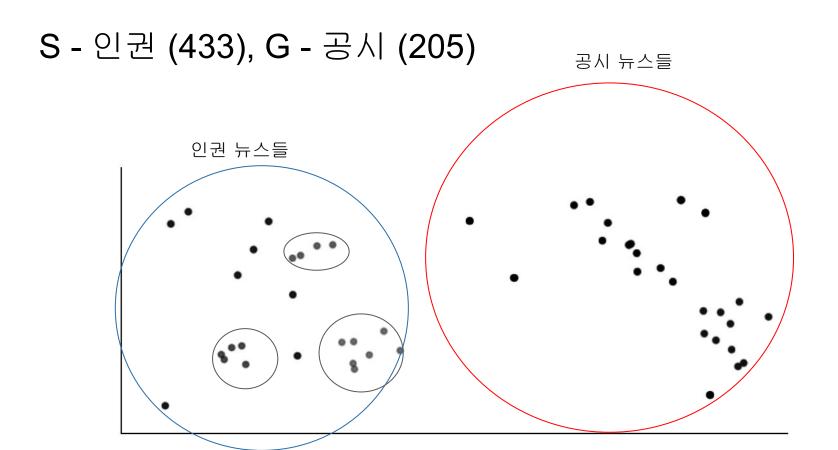
- 인권: 25개 S

총 단어 수 :

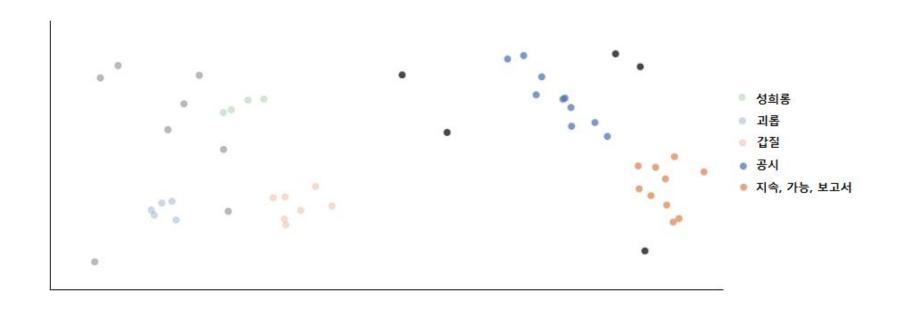
- 공시:4312개

- 인권:4006개





S - 인권 (433), G - 공시 (205)



그룹 A - 1 [인권]

공통 단어 : 갑질

뉴스 예시: 1. 한전 사내 갑질 또 터졌다김종갑 사장 입 열까 https://www.sentv.co.kr/news/view/572861

- 2. 오스템임플란트주말출근강요 등 직원갑질로시끌 http://www.wolyo.co.kr/news/articleView.html?idxno=100702
- 3. 삼성sdi 직장내괴롭힘 논란 회사는 갑질 금지 현실은 딴판 http://www.ekoreanews.co.kr/news/articleView.html?idxno=45016
- 4. 삼성sdi 직장내괴롭힘논란 회사는 갑질 금지 현실은 딴판 http://www.ekoreanews.co.kr/news/articleView.html?idxno=45016
- 5. 제일약품성희롱괴롭힘임금체불등 도 넘은 갑질 심각 http://www.newsian.co.kr/news/articleView.html?idxno=47655
 - 6. 직장 내 성희롱갑질 앱으로 공동신고한다공기업들 도입하는 리슨투미 중앙일보

https://mnews.joins.com/article/23874217?cloc=joongang-mhome-Group24

7. 내우외환 Ig전자 진단팀 갑질 처벌 청원에 채용비리 수사까지 http://hobbyen.co.kr/news/newsview.php?ncode=1065590725559672

그룹 A - 2 [인권]

공통 단어 : 괴롭

- 뉴스 예시: 1. 삼성sdi 직장내괴롭힘 논란 회사는 갑질 금지 현실은 딴판 http://www.ekoreanews.co.kr/news/articleView.html?idxno=45016
 - 2. 임원 폭행 물의 제일약품이번엔 직원 절반 직장 내 괴롭힘 당했다 https://www.ajunews.com/view/20210311163927028
 - 3. 박영사 직장 내 괴롭힘 대응 지침서 일터에서 어려움을 겪고 ... https://www.newswire.co.kr/newsRead.php?no=913401
 - 4. 한전 직장 내 괴롭힘 신고자 부당전보 의혹으로 피소 http://news.khan.co.kr/kh_news/khan_art_view.html?artid=202007020600025&am p;code=940301
 - 5. 쿠팡 물류센터 노동자 노조 밴드 가입하니 괴롭힘 시작 http://www.hani.co.kr/arti/society/society_general/992223.html

그룹 A - 3 [인권]

공통 단어 : 성희롱

- 뉴스 예시: 1. 한전kdn 성희롱 고충상담원 2차 피해 예방 교육 실시 http://www.fnnews21.com/news/articleView.html?idxno=32977
 - 2. 성희롱괴롭힘임금체불 제일약품의 불편한 민낯 http://www.sisaweek.com/news/articleView.html?idxno=142416
 - **3**. 매맞고 성희롱에 폭언 무급으로 얼룩진 재외동포재단 한상기업 ... http://www.nodongilbo.com/news/articleView.html?idxno=336378
 - 4. 막말부터 성희롱까지재택근무에도 직장내 괴롭힘 여전 https://www.asiae.co.kr/article/2020082810101863263

그룹 A - 4 [인권]

공통 단어:X

뉴스 예시: 1. 사설공공기관기업 장애인 의무고용 준수해야 http://honam.co.kr/detail/nPE2fi/636626

- 2. 제주녹색당 한국가스공사 미얀마 군부 지원 중단해야 http://www.jejusori.net/news/articleView.html?idxno=328231
- 3. 세계시민선언 포스코의 미얀마 군부배당 지급 중단 해명 반박 http://www.lcnews.co.kr/news/articleView.html?idxno=14128
- 4. e경제뉴스 모바일 사이트 30대 기업 여성 평균임금 남성의 70 http://www.e-conomy.co.kr/news/articleView.html?idxno=30399
- 5. 미얀마 사태에 내몰린 기업들 https://www.snmnews.com/news/articleView.html?idxno=477683
- 6. kb금융 2021 블룸버그 양성평등 지수 기업으로 3년 연속 선정 https://www.lawissue.co.kr/view.php?ud=202101281740396253204ead0791 12
- 7. 219개 단체 한국 정부 미얀마 민주주의 위한 조치 마련하라 http://www.newscham.net/news/view.php?board=news&category1=38&nid=105767
- 8. 분다지만 대기업 직원 성비 불균형 여전 분다지만 대기업 직원 성비 불균형 여전 https://www.fetv.co.kr/news/article.html?no=78256
- 9. 구광모 lg그룹 회장 새해 중점 경영과제 품질환경안전 http://www.mediasr.co.kr/news/articleView.html?idxno=64714

그룹 B - 1 [공시]

공통 단어 : 공시

뉴스 예시 : 1. 헬릭스미스불성실 공시법인 지정예고 http://www.mdtoday.co.kr/mdtoday/index.html?no=377986

- 2. 메디톡스 100억 규모 자사주 처분 허위공시 논란 http://news.mt.co.kr/mtyjew.php?no=2020042710310430477
- 3. 거래소 티웨이항공에 불성실공시법인지정 예고 https://www.edaily.co.kr/news/read?newsld=03696566625839768&mediaCodeNo=257
- 4. 헬릭스미스 불성실 공시 법인 지정 http://www.newspim.com/news/view/20200313000040
- 5. 두산퓨얼셀 공시불이행 불성실공시법인지정 https://www.edaily.co.kr/news/read?newsld=03876966628981352&mediaCodeNo=257
- 6. 상법 개정에 주총 대란 현실화공시누락기업 속출 https://news.naver.com/main/read.nhn?mode=LSD&mid=sec&sid1=101&oid=016&aid=0001636514
- 7. 국내 기업 사업보고서 조기제출 탓에 정정공시 대란 우려 https://www.ajunews.com/view/20210307121612825
- 8. 거래소 한국주철관공업불성실공시법인지정예고 http://www.edaily.co.kr/news/newspath.asp?newsid=04011446625803032
- 9. 금융그룹 첫 공시자본적정성교보 333내부거래 삼성 9.6조 매일경제 https://www.mk.co.kr/news/economy/view/2020/09/995127/
- 10. 네이버 정보보호 현황 공개되나공시 의무제 도입 https://news.mt.co.kr/mtview.php?no=2020101111187612155&VN

그룹 B - 2 [공시]

공통 단어 : 보고서

뉴스 예시: 1. 국내기업 사업보고서 조기제출 탓에 정정공시 대란 우려

2. 넥센타이어지속가능경영보고서 발간 http://www.shinailbo.co.kr/news/articleView.html?idxno=1316600

3. 한전 지속가능경영보고서발간 https://www.kharn.kr/news/article.html?no=14287

4. Igu 인터랙티브 지속가능경영보고서발간

https://www.g-enews.com/view.php?ud=202011130927034496becd74a0af_1&ssk=2017011301560109486_1

- 5. 한화생명 2020 지속가능경영보고서...지속가능경영 중요 이슈 ... http://www.ilyoseoul.co.kr/news/articleView.html?idxno=422990
- 6. kt 통합보고서에 지속가능경영 활동 담아 재무적 비재무적 ... http://www.cbci.co.kr/news/articleView.html?idxno=409027
- 7. 대림산업 지속가능경영 보고서 발간 배원복 사회적 책임 다한다 http://www.businesspost.co.kr/BP?command=mobile_view&num=195706
- 8. kt 2020 통합보고서 지속가능경영성과 공개 http://hyundaenews.com/93030
- 9. 롯데케미칼 지속가능경영보고서발간김교현 대표 스페셜티 ... https://www.asiatodav.co.kr/view.php?key=20200706010002827&ref=section_list
- 10. 동아쏘시오홀딩스2019 지속가능경영보고서가마솥 ... https://www.fetimes.co.kr/news/articleView.html?idxno=95367

그룹 B - 3 [공시]

공통 단어 : X

- 뉴스 예시: 1. 기업법 시행 초읽기 탄소 중립 피할 수 없는 과제 https://www.ajunews.com/view/20201210221345354
 - 2. 두산밥캣 두산 산업차량bg 인수 추진건설장비 시너지 https://news.mt.co.kr/mtview.php?no=2021022222151464046
 - 3. ff 회사 분할 결정 등 서울경제 https://sedaily.com/NewsView/1ZAHS35FRX
 - **4**. 대한제분 태영건설 한국항공우주 등 매일경제 https://www.mk.co.kr/news/stock/view/2020/12/1339822/
 - 5. 오리온 기타 경영사항 http://vip.mk.co.kr/news/view/21/28/843245.html

최종 목표 : 성능향상 -> 비용절감

솔루션 : 비지도학습 모델

기대효과

- 데이터 기반이므로 적은 샘플 수로 분류 가능.
- 새로운 카테고리를 생성시 객관적 기준 마련(데이터 및 알고리즘 기반)
- 잘못된 레이블링 자동 수정 가능

GPT1,2,3 는 유료 버전으로 상용화되어있는 상태. GPT3.5 는 ChatGPT로 불리며 베타버전인 상태. 따라서 ChatGPT기반의 상용화된 서비스는 현재 없다.

특징:

기존 GPT3 버전까지는 학습용 데이터를 대량으로 사용 + 미세조정 의 방법으로 만들었다. 이번 GPT3.5(ChatGPT)의 경우는 강화학습을 기반으로 훈련을 하였다.

투자 관련 :

Ms가 openAl chatgpt에 12조 투자

생성AI관련 투자:

프로플루언트: 110억원 / 단백질 생성 Al

슈퍼노멀: 123억원 / 회의 내용을 자동으로 정리해주는 서비스 (GPT3기반)

생성AI 서비스(상용 서비스는 현재 GPT3 사용 중):

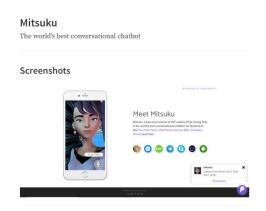
- 스지키 쇼조텐, 앳워터 브루어리 / 술 제조에 활용
- 줌 / 대화형 AI 솔루션 , 채팅, 컨택센터 플랫폼 등과 통합하는 것
- 이크림 / 스토리 창작
- 뤼튼테크놀로지스 / 광고 문구, 글 초안 제작
- 미국 매사추세츠주 상원의원과 하원의원이 '챗GPT'를 활용해 인공지능(AI) 기술을 규제하는 법안(초안)을 작성

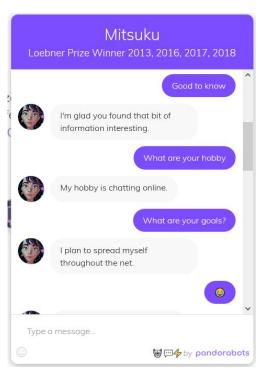
Chatgpt 기반의 서비스 조사:

- 라이센스 문제로 공식적으로 서비스화된 사례는 안나온다.

대표 상용 서비스

- OpenAl's API
- Replika AI : 가상 AI 챗 아바타 서비스
- Mitsuku AI : 가상 AI 챗 아바타 서비스
- Hugging Face's Transformer





ESG Chatbot

회사가 물 관련 위험에 대한 노출을 평가하는 것에 대한 예시를 보여줘



다음은 회사가 물 관련 위험에 대한 노출을 어떻게 평가할 수 있는지에 대한 예입니다.

- 1. 물 위험 평가 수행: 회사는 물리적, 규제 및 평판 위험을 포함하여 운영과 관련된 잠재적 인 물관련 위험을 식별하기 위해 물 위험 평가를 수행 할 수 있습니다.여기에는 공급망, 생산 공정 및 폐기물 관리 관현과 관련된 물관련 위험을 분석하는 것이 포함될 수 있습니다.
- 2. 물 발자국 분석 : 회사는 운영에서 회사가 사용하는 총 물량을 측정하여 전체 물 사용량과 지역 수자원 자원에 미치는 잠재적 영향을 결정하기 위해 물 발자국을 분석 할 수 있습니다.
- 3. 이해 관계자와의 관계 : 회사는 지역 사회, 정부 기관 및 환경 단체를 포함한 이해 관계자와 교 류하여 운영과 관련된 물 관련 위험을 더 잘 이해하고 개선 기회를 식별 할 수 있습니다.
- 4. 물 관리 관행 평가 : 회사는 물 사용의 효율성 및 물 절약에 사용하는 방법과 같은 물 관리 관행을 평가하고, 물 관련 위험에 대한 잠재적 노출을 결정하고 개선을위한 영역을 식별 할 수 있습니다.
- 5. 수자원 규정 준수 평가 : 회사는 규제 위험에 대한 잠재적 노출을 결정하기 위해 물 사용 제한 및 배출 제한과 같은 지역 및 국가 수자원 규정 준수를 평가할 수 있습니다.
- 6. 도구 및 프레임 워크 사용: 회사는 물 위험 평가 및 지속 가능성 프레임 워크와 같은 도구 및 프레임 워크를 사용하여 물 관련 위험에 대한 노출에 대한 더 깊은 이해를 얻고 개선 기회를 식별할 수 있습니다.
- 7. 정기 모니터링 및 검토 : 회사는 이러한 위험을 관리하는 데 적극적으로 보장하기 위해 수자원 사용, 수자원 관리 관행 및 수자원 규정 준수를 포함하여 물 관련 위험에 대한 노출을 정기적으로 모니터링하고 검토 할 수 있습니다.

이 예는 회사가 내부 및 외부 요인을 모두 고려하고 다양한 도구와 기술을 사용하여 이러한 위험에 대한 포괄적 인 이해를 얻는 방법을 다각적으로 사용하여 물 관련 위험에 대한 노출을 평가할 수있는 방법을 보여줍니다.그렇게함으로써 회사는 이러한 위험에 대한 노출을 줄이고보다 지속 가능한 미래를 보장하기 위해 조치를 취할 수 있습니다.