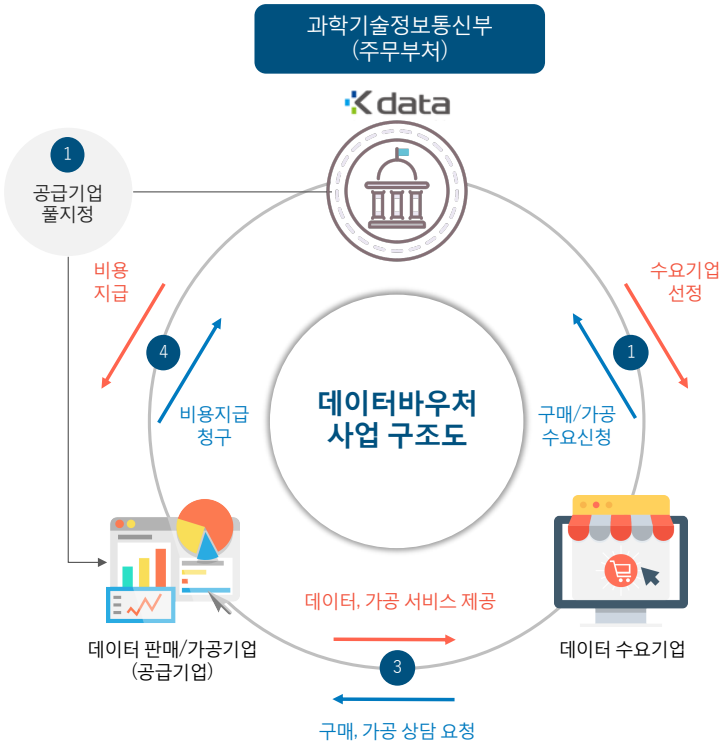
과학기술정보통신부 · 한국데이터산업진흥원

2021년 데이터바우처 지원사업 안내 및 기업 소개

XINAPSE

데이터바우처지원사업으로 우리 기업이 보유한 데이터를 분석·가공할 수 있습니다.

데이터바우처지원사업은 중소·벤처 기업, 소상공인 등의
데이터 활용을 촉진하기 위해 데이터 전처리, 태깅, 레이블링, 분석 등
데이터 가공 서비스 비용을 지원하는 한국데이터산업진흥원 사업입니다.



지원 자격은 어떻게 되나요?

데이터를 활용하여 혁신적인 서비스를 창출할 수 있는
국내 중소기업, 스타트업, 1인 창조기업이 지원할 수 있습니다.

어떤 지원을 받을 수 있나요?

심사를 거쳐 선발된 프로젝트에 대해 최대 **일반가공 4천 5백만원,**
AI가공 7천만원 상당의 바우처를 지급합니다.(기업부담금 25%, 전액 현물가능)

어떻게 지원해야 하나요?

수요기업은 사업에 지원하기 전에 **반드시 가공기업과 업무범위에**
대한 사전 협의 후 수행계획서와 함께 사전협의서를 제출해야 합니다.
이후 진행되는 발표평가 역시 가공기업과 함께 참여할 수 있습니다.

자이냅스는
데이터바우처지원사업의
공식 가공기업입니다.

다량의 데이터 전처리, 레이블링, 토픽 모델링, 음성합성(TTS)과 같은
자이냅스의 데이터 분석/가공 기술은 다양한 분야에서 활용되고 있습니다.

KIET 산업연구원
뉴스 데이터를 활용한 제조업 주요 업종의
경기동향 조사(자동차, 반도체)

KIET 산업연구원
뉴스 데이터를 활용한 제조업 주요 업종의
경기동향 조사(조선, 철강, 디스플레이 등 7개)

글로벌 투자 파트너 – MIRAE ASSET
인공지능 기반 직무별 뉴스 큐레이션 서비스 구축

세마컨설팅
감성분석과 토픽모델링을 활용한
병·의원 고객후기 분석

Platfarm
감정 데이터 레이블링을 통한 이모티콘 생성 및
매칭

MEDIAI
Medication Actionable Information
연구자 주도 대규모 감염병 백신 개발을 위한
글로벌 임상시험 데이터 가공

KDB산업은행
DT사업 비정형 분석

한국수출입은행
THE EXPORT-IMPORT BANK OF KOREA
조기경보모형 개선·구축 프로젝트

광고대행사 중개 플랫폼
MADAHM
빅데이터 기반 매칭 알고리즘을 활용한
광고대행사 매칭 플랫폼

DFINE
e-커머스 큐레이션 서비스를 위한 소비자
행동데이터 및 상품 키워드 모델링 가공/분석

자이냅스는 비정형 텍스트에 대한 전처리부터, 레이블링, 토픽모델링을 제공합니다.

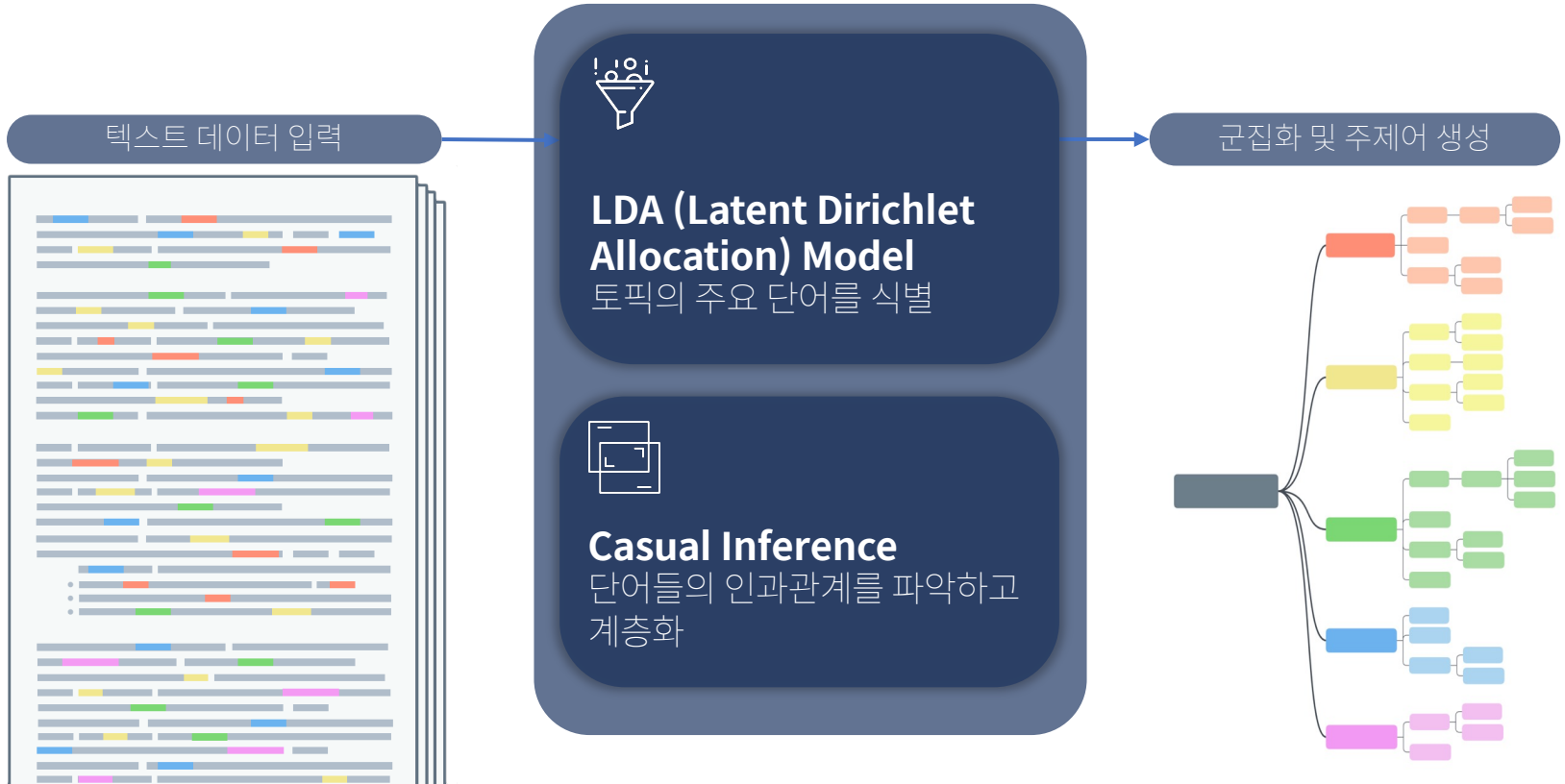
비정형 텍스트 데이터 전처리

텍스트 데이터 레이블링

데이터 군집화 및 토픽 모델링

TOPIC MODELING

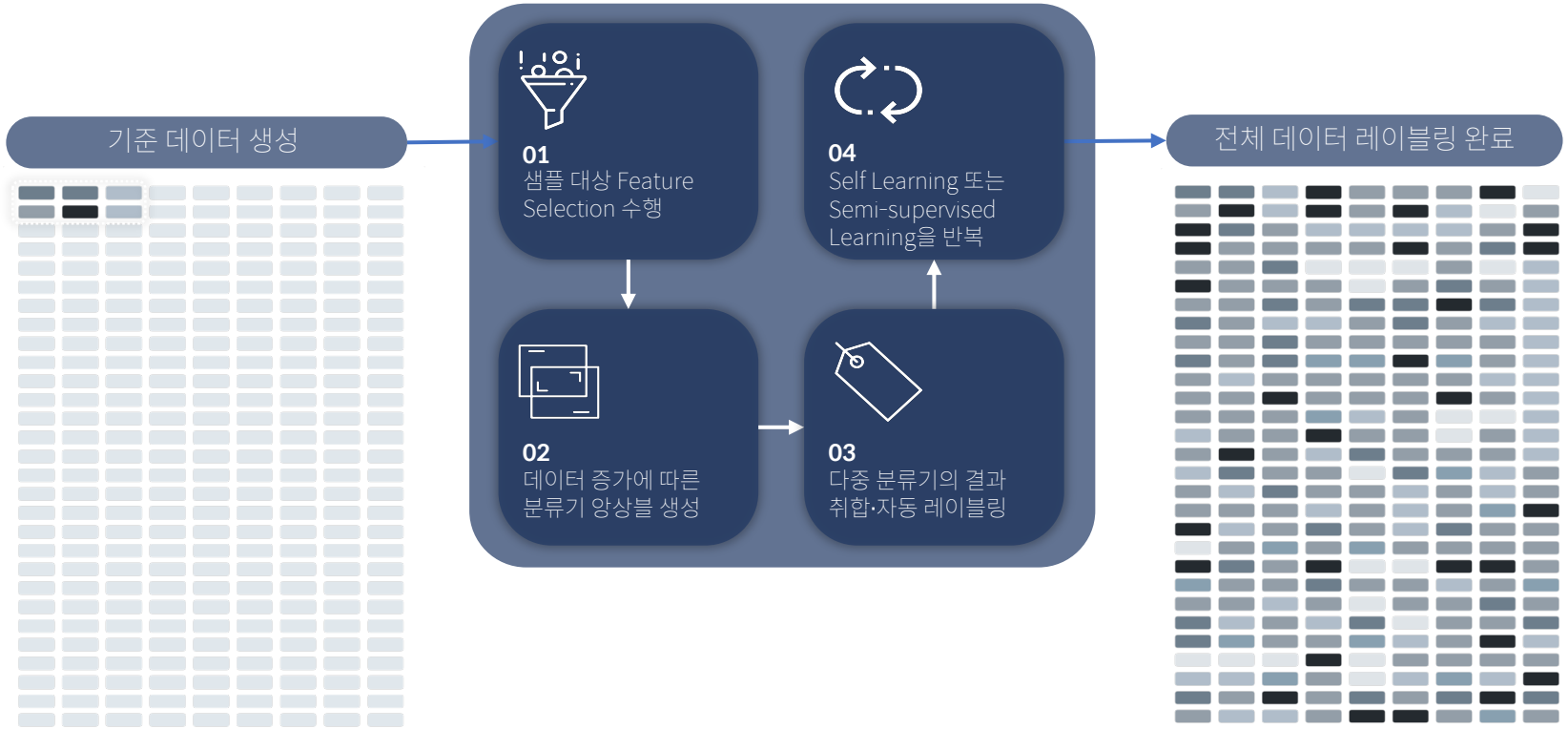
텍스트 내 유의미한 관계를 파악하고 군집화 하여 대표적인 주제어를 생성하는 기술로,
LDA(Latent Dirichlet Allocation) 모델을 통해 데이터의 확률 분포도를 계산하여 인과관계를 파악 할 수 있는 계층적 데이터 분석에 활용됨



AUTO LABELING

수기로 처리하던 기존의 데이터 분류 작업을 인공지능 기술을 통해 다량의 데이터를 빠르게 분류하는 기술로,
오토 레이블링에 필요한 학습 데이터에 따라 Self Learning 또는 Semi-supervised Learning(준지도 학습)으로 진행함

산재되어 있는 다량의 텍스트 데이터를 객관적인 기준에 맞추어 분류하여 데이터 표준화 및 일원화 분석이 가능함



데이터 전처리
(Data Preprocessing)

(주)메디아이플러스

다양한 분석을 하기 위한 첫 단계로, 수집된 데이터의 단위, 표시형식 등을 표준화 하고 정규화 하는 처리 과정임

국내외 산재되어 있는 임상 관련 데이터를 쉽게 검색할 수 있도록 총 194만건(국내 1,600여건, 해외 192만여건)의 임상시험 데이터를 요구 조건과 기준에 맞추어 전처리 작업을 수행

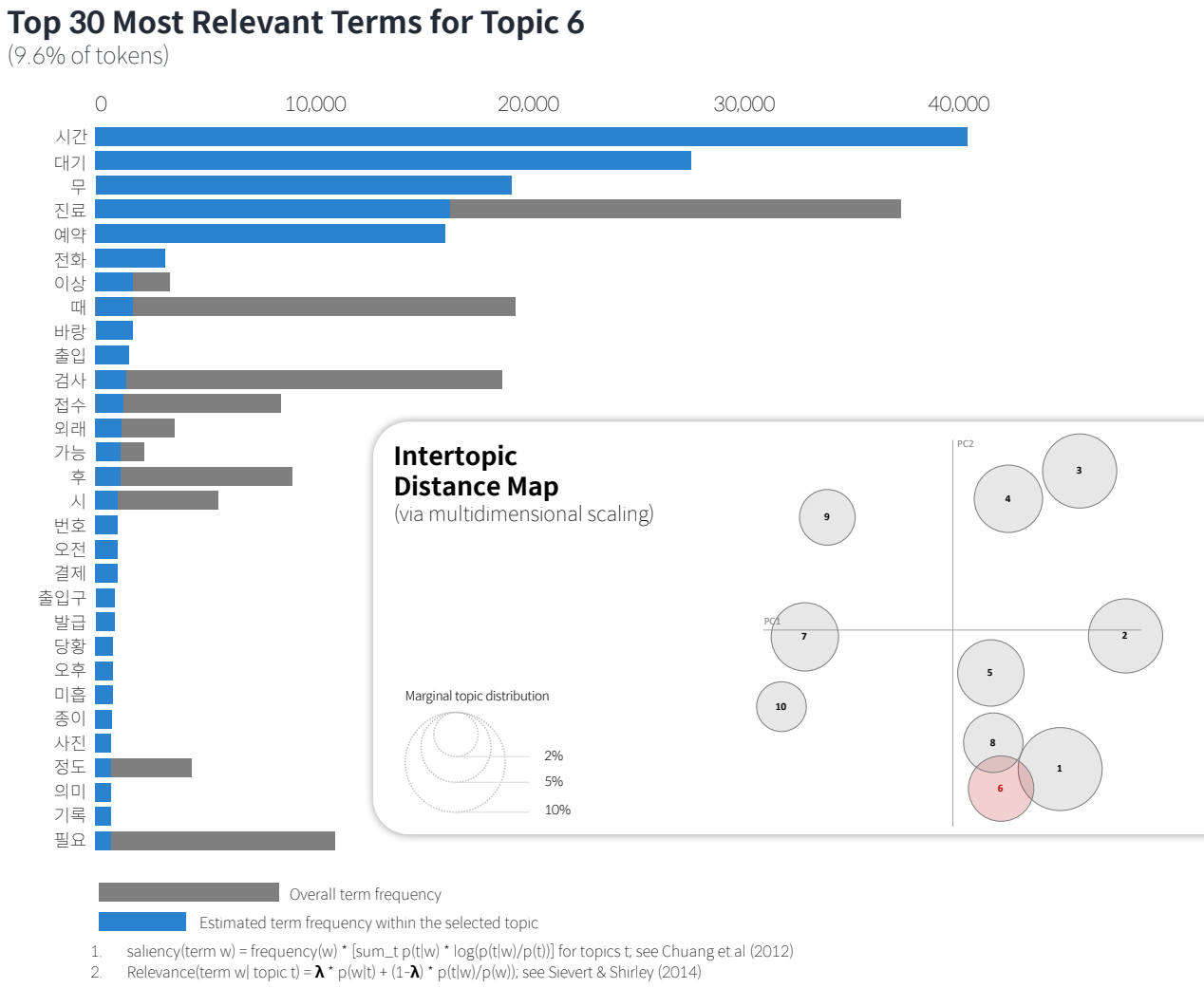
	kct_id	scientific_name_en	scientific_class_en	affiliation_en	zip_code	affiliation_tel	public_name_en	public_scientific_class_en	public_affiliation_en
0	KCT0000001		MD	University of Ulsan	44610			CRC	Asan Medical Center
1	KCT0000002		MD	National Cancer Center	10408			MD	National Cancer Center
2	KCT0000003		MD, PhD	Seoul National University Bundang Hospital	13620			MD, MS	Seoul National University B
3	KCT0000004		MD	Asan Medical Center	5505			MD	Asan Medical Center
4	KCT0000005		Professor, MD	The Catholic University of Korea, Seoul St. Mary's Hospital	6591			Clinical professor	The Catholic University of K
5	KCT0000006		MD	Chonbuk National University Hospital	54907	063-250-2185		College	Chonbuk National Universit
6	KCT0000007		professor	The Catholic University of Korea, Seoul St. Mary's Hospital	6591			Instructor	The Catholic University of K
7	KCT0000008		MD	Seoul National University Hospital	3080			Master	JW CreaGene
8	KCT0000009		PhD	Chonbuk National University Hospital	54907	063-250-2185		college	Chonbuk National Universit
9	KCT0000010		M.D., Ph.D.	Seoul National University	8826			M.D.	Seoul National University
10	KCT0000011		MD, PhD	Yonsei University, Wonju Severance Christian Hospital	26426			MD, PhD	Ewha Womans University
11	KCT0000012		Professor, M.D.	The Catholic University of Korea, Seoul St. Mary's Hospital	6591			Clinical professor	The Catholic University of K
12	KCT0000013		MD	Pusan National University Yangsan Hospital	50612			nurse	Pusan National University Y
13	KCT0000014		MD	Kyungpook National University Hospital	41944			MD	Kyungpook National Univer
14	KCT0000015		MD	Asan Medical Center	5505			MD	Asan Medical Center
15	KCT0000017		MD	Kyungpook National University Hospital	41944			MD	Kyungpook National Univer
16	KCT0000018		MD,PhD	Seoul National University Bundang Hospital	13620			MD	Seoul National University B
17	KCT0000019		KMD	Kyung Hee University	2453			KMD	Kyung Hee University
18	KCT0000020		MD	Pusan National University Yangsan Hospital	50612			MD	Pusan National University Y
19	KCT0000021		MD, PhD	Konkuk University Medical Center	5030			MD, PhD	Konkuk University Medical
20	KCT0000022		professor	Seoul National University Hospital	3080			sub-investigator	Seoul National University H
21	KCT0000023		MD	Asan Medical Center	5505			MD	Asan Medical Center
22	KCT0000024		MD, Ph D	Asan Medical Center	5505			MD	Asan Medical Center
23	KCT0000025		MD, PhD	Yonsei University Health System, Gangnam Severance Hospital	6273			MD, PhD	Yonsei University Health Sy
24	KCT0000026		MD	Asan Medical Center	5505			MD	Asan Medical Center
25	KCT0000027		OMD	Dongguk University	4620			OMD	Dongguk University
26	KCT0000028		MD	Asan Medical Center	5505			MD	Asan Medical Center
27	KCT0000029		MD	Asan Medical Center	5505			MD	Asan Medical Center
28	KCT0000030		MD,PhD	Koera University Guro Hospital	8308			MD,PhD	Koera University Guro Hosp
29	KCT0000031		MD	Hanyang University Seoul Hospital	4763			MD	Hanyang University
30	KCT0000032		MD	Kangbuk Samsung Medical Center	3181			MD	Sungkyunkwan University
31	KCT0000033		MD, PhD	Seoul National University Hospital	3080			MD	Seoul National University H
32	KCT0000034		MD, PhD	Koera University Guro Hospital	8308			MD, PhD	Koera University Guro Hosp
33	KCT0000035		MD	Hanyang University Seoul Hospital	4763			MD	Hanyang University Seoul H
34	KCT0000036		MD, Ph D	Kyungpook National University Hospital	41944			MD	Kyungpook National Univer
35	KCT0000037		MD	Seoul National University Hospital	3080			MD	Seoul National University H
36	KCT0000038		professor	The Catholic University of Korea, Seoul St. Mary's Hospital	6591			assistant professor	The Catholic University of K
37	KCT0000039		MD	Keimyung University	42403 42601			MD	Keimyung University
38	KCT0000041		M.D.	Yonsei University Health System, Severance Hospital	3722			M.D.	Yonsei University Health Sy
39	KCT0000042		MD, Ph.D	Hanyang University Guri Hospital	11923			MD, Ph.D	Hanyang University Seoul H
40	KCT0000043		MD, PhD	Seoul National University Hospital	3080			MD	Seoul National University H
41	KCT0000044		MD, PhD	Koera University Guro Hospital	8308			MD, PhD	Koera University Guro Hosp
42	KCT0000045		MD	Keimyung University	42403 42601			MD	Keimyung University

고객후기(Voice of Customer) 분석

(주)세마그룹

자이냅스의 고객 후기 분석은 도메인에 구애 받지 않고
고객 후기 및 댓글 분석이 가능하도록 모델을 개발함

병의원의 고객 후기 데이터 24만건의 데이터를 병원별/시간추이별로
토픽별 감성분석을 통해 유의미한 결과를 도출함

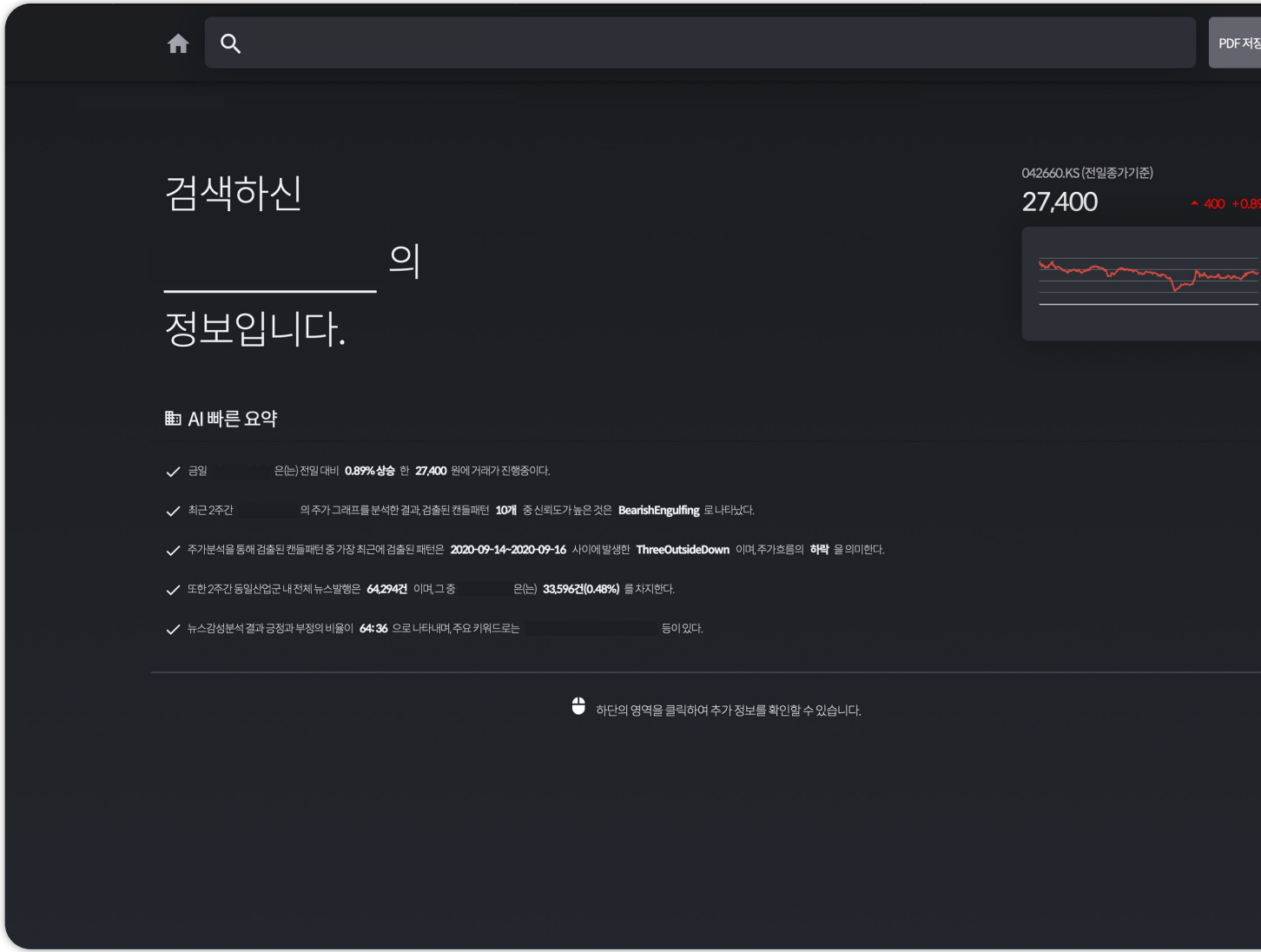


뉴스데이터를 활용한 산업 동향 분석

산업연구원

뉴스데이터를 사용하여 비정형 데이터를 정량화 하고
주제에 맞게 토픽모델링, 데이터 레이블링 등 다양한 기술이 사용됨

분석하고자 하는 산업의 뉴스 약 80만건(11년치)을 투입하여
산업의 변화를 파악하고, 시장의 변화와 기업 비즈니스에 활용됨

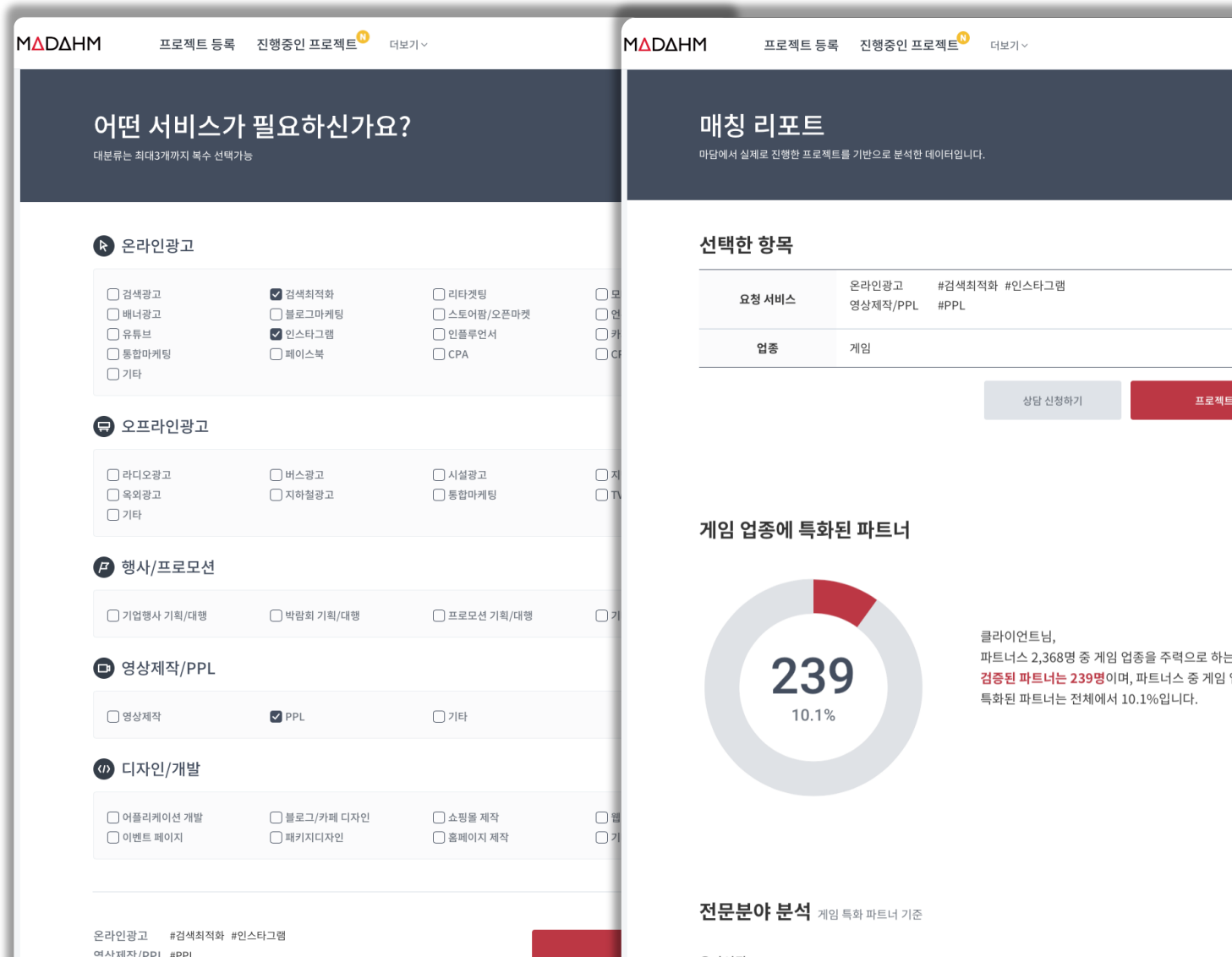


기업 매칭을 위한 알고리즘 개발

(주)마담

서비스 기획에 따라 데이터 간 유사도 및 가중치를 반영한
매칭 알고리즘을 개발함

광고를 필요로 하는 기업이 자신의 광고 목적에 맞는
최적의 대행사를 찾을 수 있도록 알고리즘이 구현되었으며,
도메인에 따른 튜닝이 가능



이모티콘 매칭을 위한 오토레이블링 기술 적용

(주) 플랫팜

(주)플랫팜은 텍스트를 입력하면 이모티콘을 추천해주는 모히톡(mojitok) 서비스를 제공하는 기업으로 기존에는 텍스트와 이모티콘을 매칭하기 위하여 수행하는 레이블링을 사람이 일일이 수기로 작업하였으나, 자이냅스 오토레이블링을 적용하여 효율성을 개선하였습니다.

소비자 행동 데이터 기반 상품 추천 알고리즘 설계

(주) 디파인

국내 E-커머스 시장의 소비자 행동패턴 데이터를 바탕으로 각 해당 도메인 로그 데이터, 구매/취소 등 상품 데이터를 분석하여 기업의 온라인 시장을 위한 의사결정 근거자료로 활용(URL 데이터 약 4,000만건)하였습니다.

부실징후 사전인지 모델 수립

한국수출입은행·KDB산업은행

은행으로부터 대출받은 기업들의 리스크를 정형 데이터로 파악하는 방식은 적시성이 떨어지므로, 이에 대한 리스크 관리를 비정형 빅데이터를 통해 파악하고자 하였습니다.

Thank you.

XINAPSE

데이터팀 김대웅 팀장

(06179) 서울특별시 강남구 테헤란로 86길 14, 윤천빌딩 8층

Web. www.xinapse.ai

Email. dwkim@xinapse.ai

Phone. 02-6052-5611, 5587