

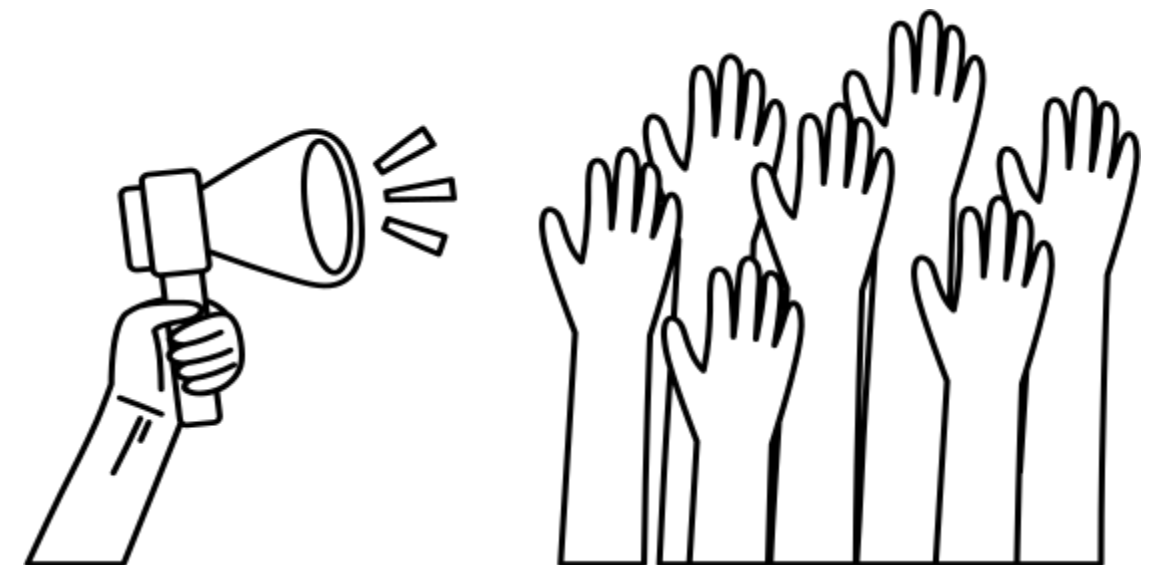
뉴스 기사 기반

진보 보수 지수 측정을 통한

# 정치적 편향성 완화 프로젝트

34-35기 NLP팀 주제분석

김현우 박윤서 배소정 정재훈



CONTENTS

# 소개 목차

---

01

문제 배경

02

프로세스 요약

03

활용 데이터

04

국회 회의록 편향성  
유의성 검정

05

정치 편향 지수 산출

06

편향성 완화 모델

---

01

# 문제 배경

# 01 문제 배경

## 정파적 지표화

### 정파적 지표화 (Partisan Indexing)

언론이 자신의 **정치적 성향**에 따라 전문가 인용이나 사실 보도에 **편향**을 드러내는 보도 행태 혹은 현상

참고 문헌 : 이종혁(2022), 보수 언론과 진보 언론의 북한 전문가 활용 방식의 차이 탐색

#### 李 유죄 파기환송, 조선 “대선 출마 자격 상실” 한겨레 “사법부 개입”

[아침신문 속아보기] 중앙일보 “유죄 예약 후보와 계엄 초래한 정당 후보 중 차악 따져야”  
사퇴한 한덕수 권한대행, 한국일보 “심판이 선수로 등판”  
韓 사퇴→崔 사퇴→李 대행에 동아일보 “국정이 이리 가법나”



박서연 기자 psynism@mediatoday.co.kr  
입력 2025.05.02 07:42

N 기자구독

후원

TALK f t g v 가가

사회

#### 윤석열 정부 다룬 사실 키워드, 조선일보는 '문재인' 한겨레는 '우려'

전 정권 비교하며 정부 칭찬한 조선-우호적 사실 한 건도 없는 한겨레...극명한 논조 차



윤수현 기자 melancholy@mediatoday.co.kr  
입력 2023.04.11 20:30

N 기자구독

후원

TALK f t g v 가가

언론사마다 같은 사건을 다루더라도 표현 방식, 제목 구성, 강조점이 다름

# 01 문제 배경

## 정치적 양극화

2011년 서울시장선거 지지 의견 모델 로지스틱 회귀분석 결과

미디어 이용	나경원 후보 지지	박원순 후보 지지
	회귀계수	회귀계수
진보신문	-.42 **	.47 ***
보수신문	.22 *	-.20
진보인터넷	-.43 **	.38 *
보수인터넷	.21	-.13

참고 문헌 : 손영준·이완수(2013), 후보 지지와 선택에 영향을 미치는 요인에 관한 연구

많은 학자들이 경험적 연구를 통해  
정파적 미디어에 대한 노출이 정치적 양극화에 영향을 미친다는 것을 입증하고 있음

참고 문헌 : Arceneaux·Johnson(2013), Changing Minds or Changing channels?

# 01 문제 배경

## 정치적 양극화

2011년 서울시장선거 지지 의견 모델 로지스틱 회귀분석 결과

미디어 이용	나경원 지지	박원순 후보 지지
	회귀계수	회귀계수
진보신문	-.42 **	.47 ***
보수신문	.22 *	-.20
보수인터넷	.21	-.13

정치적 편향을 완화하여

뉴스 회피 현상과 정치적 양극화를 방지할 수 있는 방법은 없을까?

!! 정치적 편향 정도를 측정하고 완화시킬 수 있는 연구를 진행해보자!

많은 학자들이 경험적 연구를 통해  
정파적 미디어에 대한 노출이 정치적 양극화에 영향을 미친다는 것을 입증하고 있음

참고 문헌 : Arceneaux·Johnson(2013), Changing Minds or Changing channels?

## 01 문제 배경

### 연구 필요성

#### 기존 연구의 한계점



**특정 주제**(북한, 코로나 등)에 대한 정치적 편향성 **분류 모델 개발 중심**의 연구가 많음



국내 연구들은 주로 사전에 **진보/보수 언론을 구별**하고 이를 이용하여 모델을 학습시킴



시사점만 제공하고 실질적인 정치적 편향성 **해소 방안**을 제시하지 않음



**국회 회의록**을 활용하여 정치적 편향 지수를 측정하고  
**모든 신문 기사**에 적용 가능한 **정치적 편향성 완화 모델**을 만들어보자!

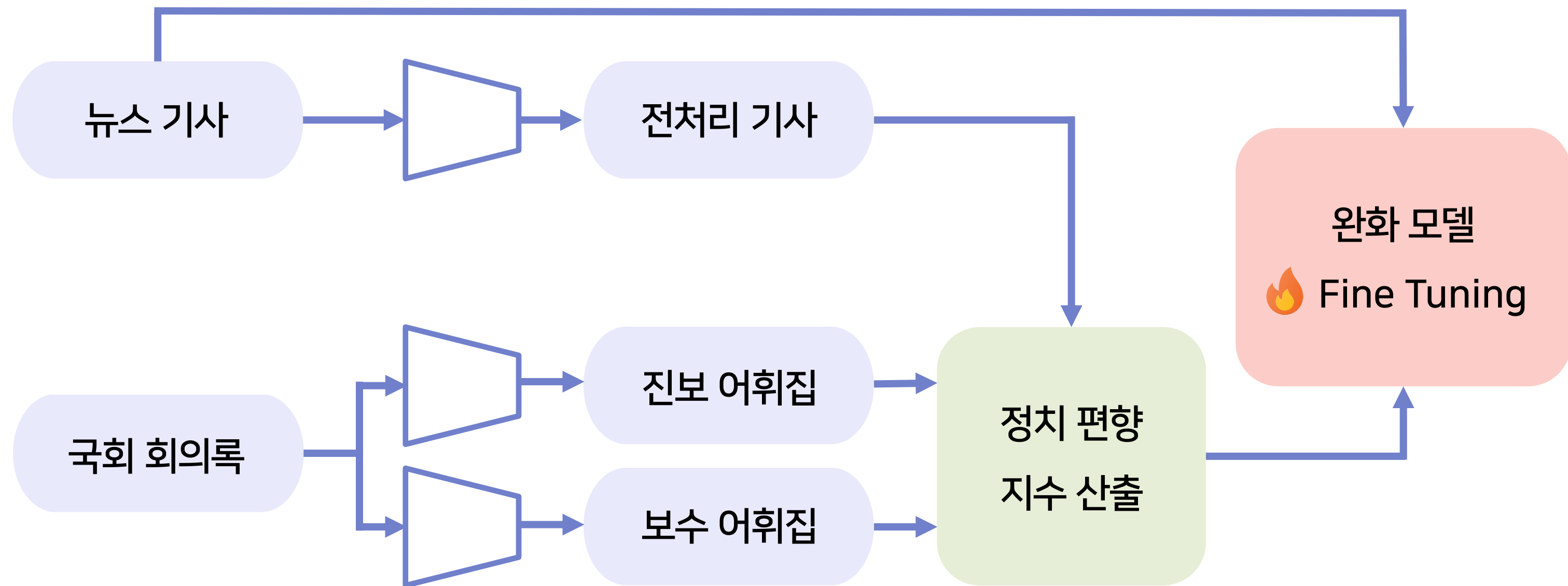
02

# 프로세스 요약



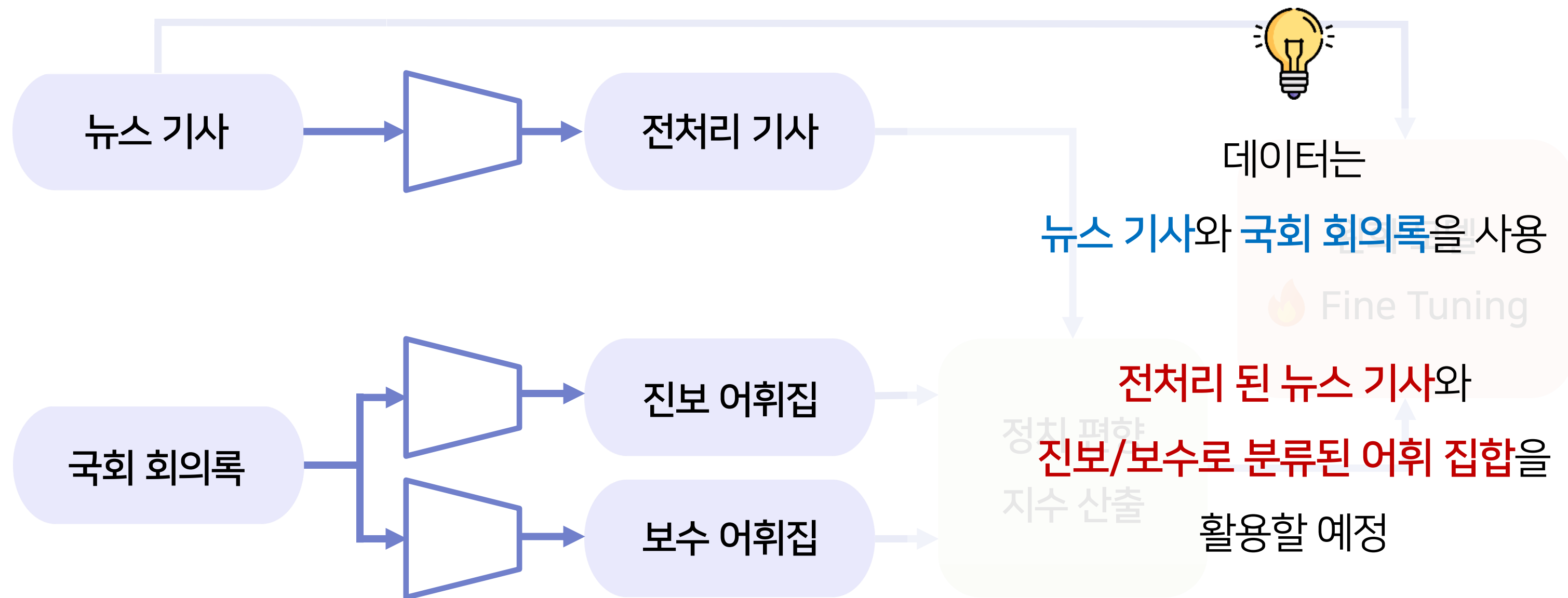
## 02 프로세스 요약

### 프로세스 흐름도



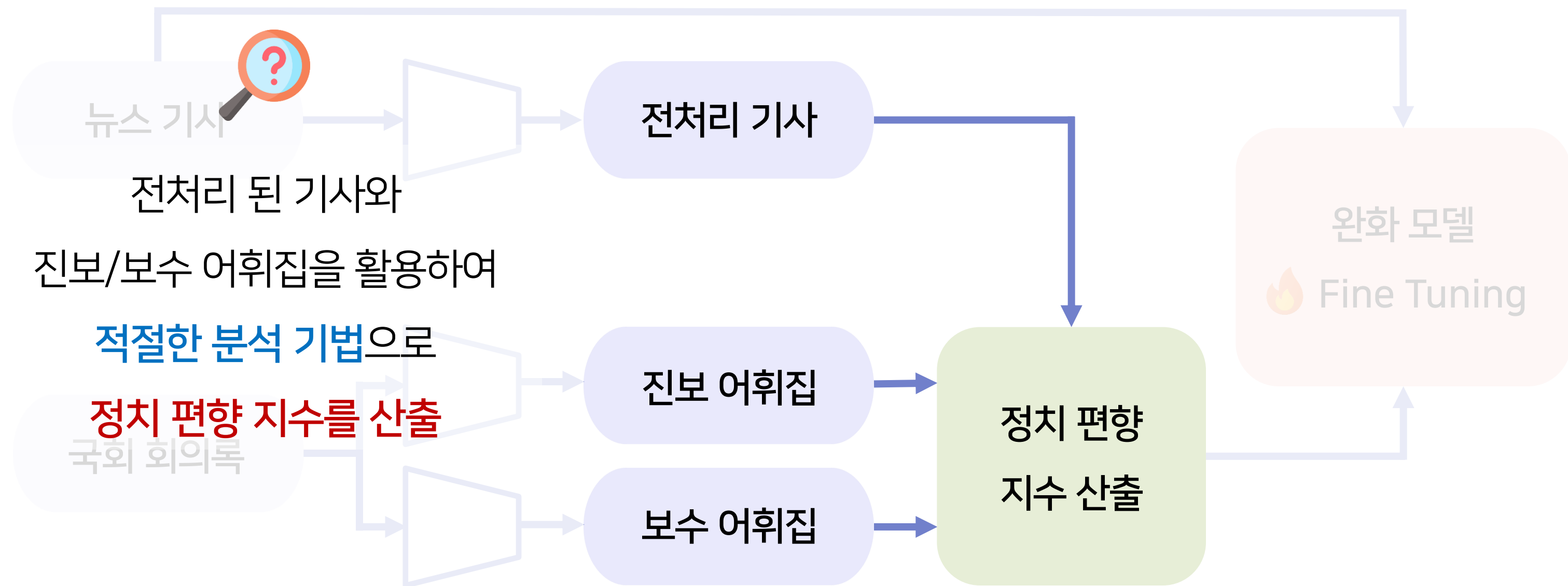
## 02 프로세스 요약

### 프로세스 흐름도



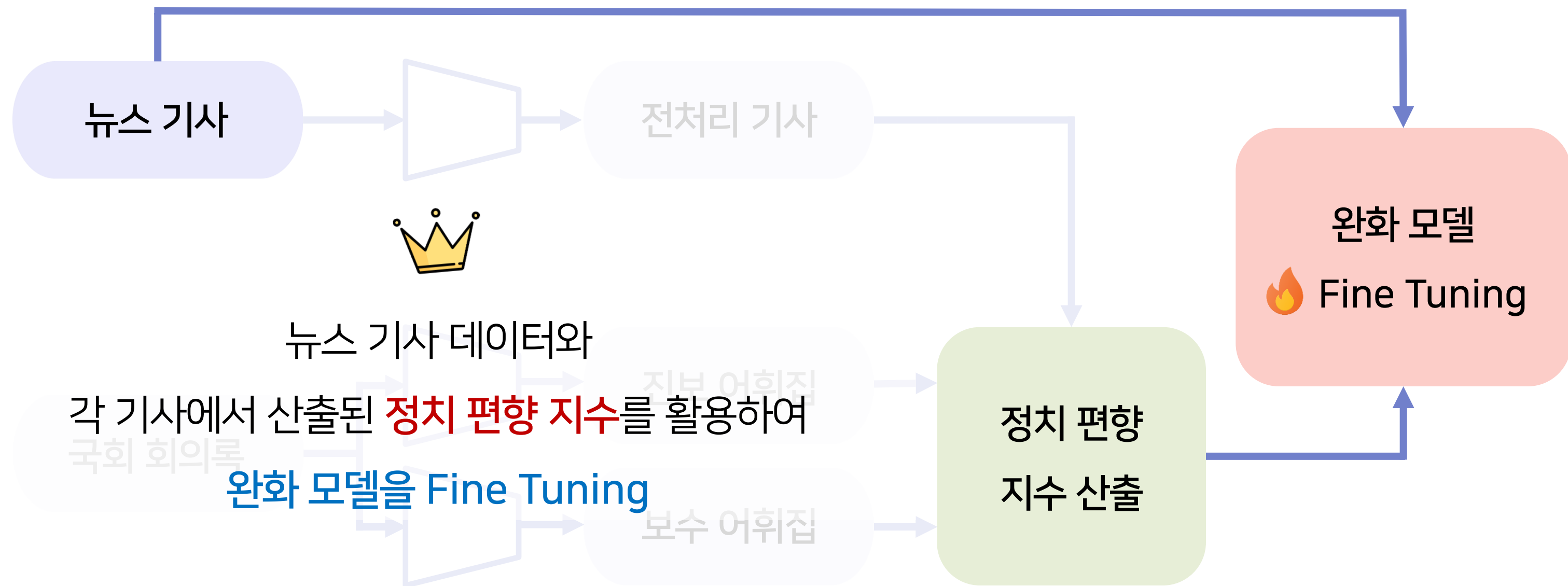
## 02 프로세스 요약

### 프로세스 흐름도



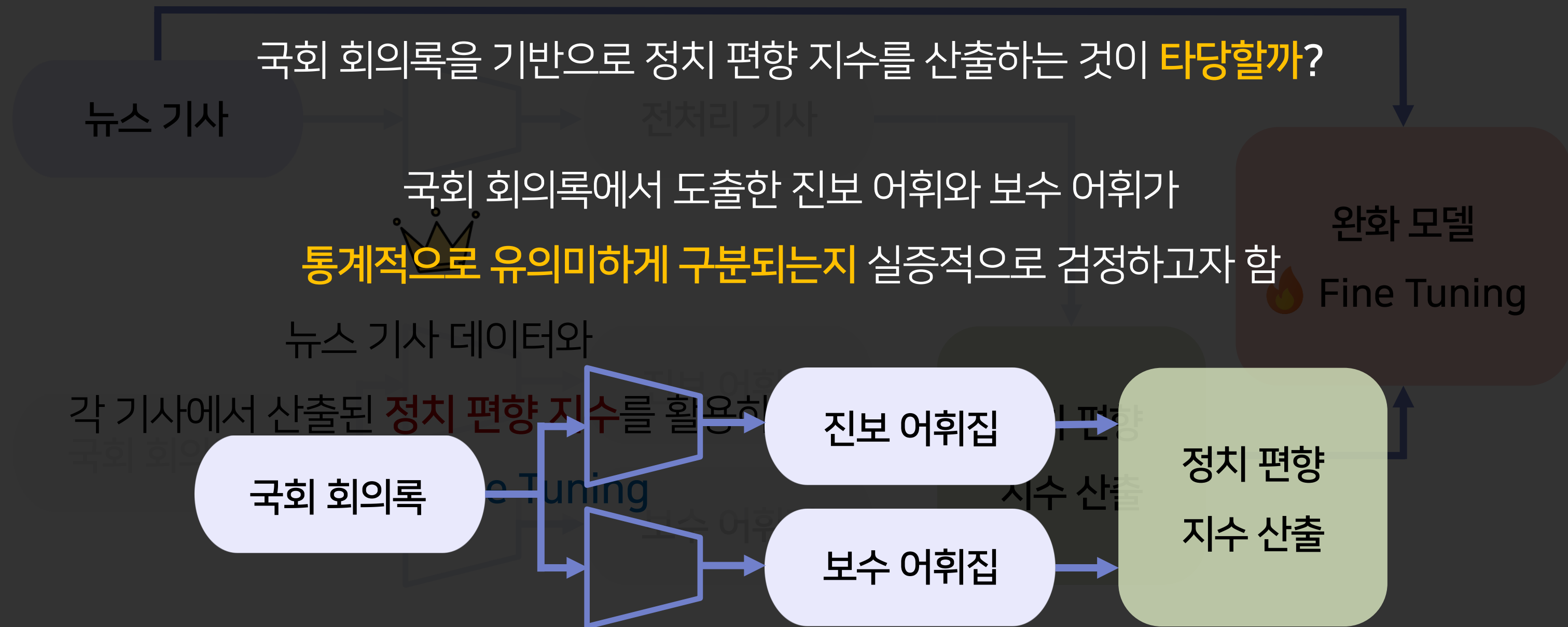
## 02 프로세스 요약

### 프로세스 흐름도



## 02 프로세스 요약

### 프로세스 흐름도



03

# 활용 데이터

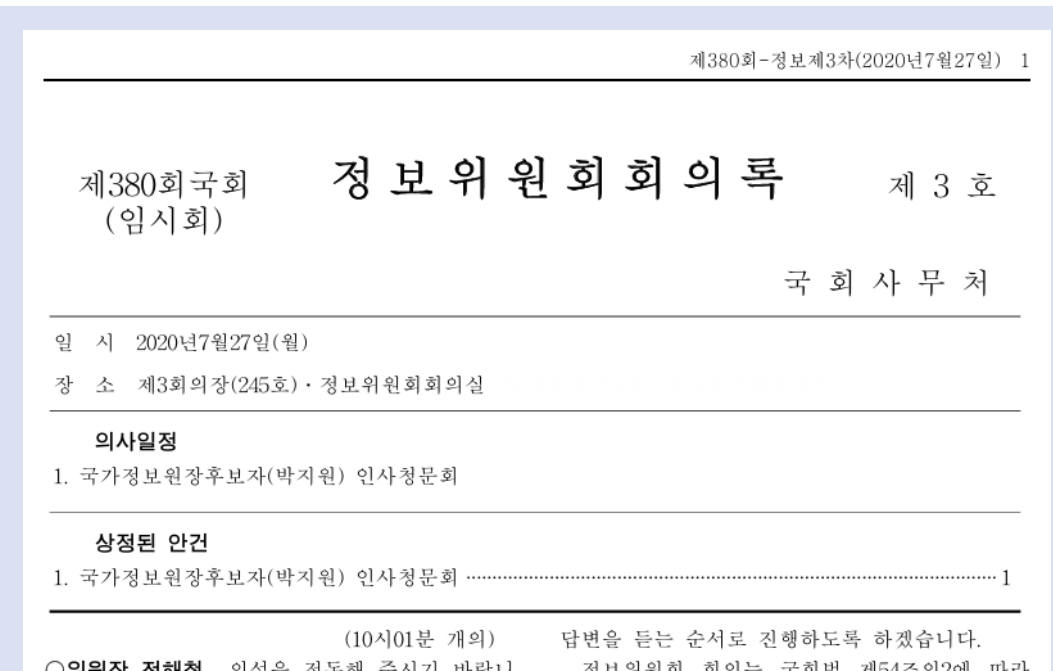
# 03 활용 데이터

## 활용 데이터



### 뉴스 기사 데이터

설정한 기간 내 사용률 상위 10개  
언론사에서 보도한 기사 데이터 수집



### 국회 회의록 데이터

설정 기간 내  
국회 회의록 데이터 수집



### 기타 데이터

국회의원 명단, 연설문 등  
분류 및 검증에 사용할 부수적인 데이터

# 03 활용 데이터

## 뉴스 기사 데이터 | 언론사 선정

공식 홈페이지 / 앱 뉴스 이용률 상위 10개 언론사를 대상으로 선정하여 기사 수집을 진행함

### 2024 언론수용자 조사

〈표 38〉 언론사 공식 홈페이지/앱 뉴스 이용률(복수응답, base=언론사 홈페이지 또는 앱 뉴스 이용자)  
문36) 지난 1주일 동안 단말기(스마트폰, 태블릿 PC, 데스크톱, 노트북 등)에 상관없이 어떤 언론사의 사이트/앱에 직접 접속하여 뉴스를 보셨습니까? [복수 응답]

구분		사려수	YTN	MBC	조선일보	JTBC	TV조선	KBS	SBS	연합뉴스 TV	중앙일보	동아일보
전체		(640)	16.1	13.8	13.7	11.8	10.3	9.3	7.4	7.3	6.9	6.9
성별	남성	(389)	15.6	13.5	14.6	11.3	9.9	8.5	6.0	7.9	5.3	7.4
	여성	(251)	16.8	14.3	12.4	12.5	10.9	10.5	9.5	6.4	9.4	6.1
연령대	19-29세	(122)	14.5	12.7	11.9	17.2	8.1	11.4	6.8	7.0	9.7	3.3
	30-39세	(140)	19.7	18.7	21.1	15.0	9.8	8.2	5.8	6.7	4.8	3.4
	40-49세	(152)	17.8	9.6	9.3	12.4	9.8	7.4	12.0	8.1	8.0	7.6
	50-59세	(131)	16.5	16.9	17.2	5.1	12.9	9.0	6.3	9.3	5.9	12.2
	60-69세	(69)	7.3	10.0	10.3	9.0	12.2	11.8	5.2	6.8	5.6	4.3
	70대 이상	(26)	15.9	12.5	0.0	5.6	8.0	11.3	2.3	0.0	7.0	18.6

한국언론진흥재단 2024 언론 수용자 조사 결과



YTN, MBC, 조선일보, JTBC,  
TV조선, KBS, SBS, 연합뉴스TV,  
중앙일보, 동아일보가 포함됨을 확인



## 03 활용 데이터

### 뉴스 기사 데이터 | 기간 선정

정치적으로 큰 이슈였던 윤석열 전 대통령의 탄핵 선고일인  
2025년 4월 4일을 기준으로 총 5주 간으로 기간 선정

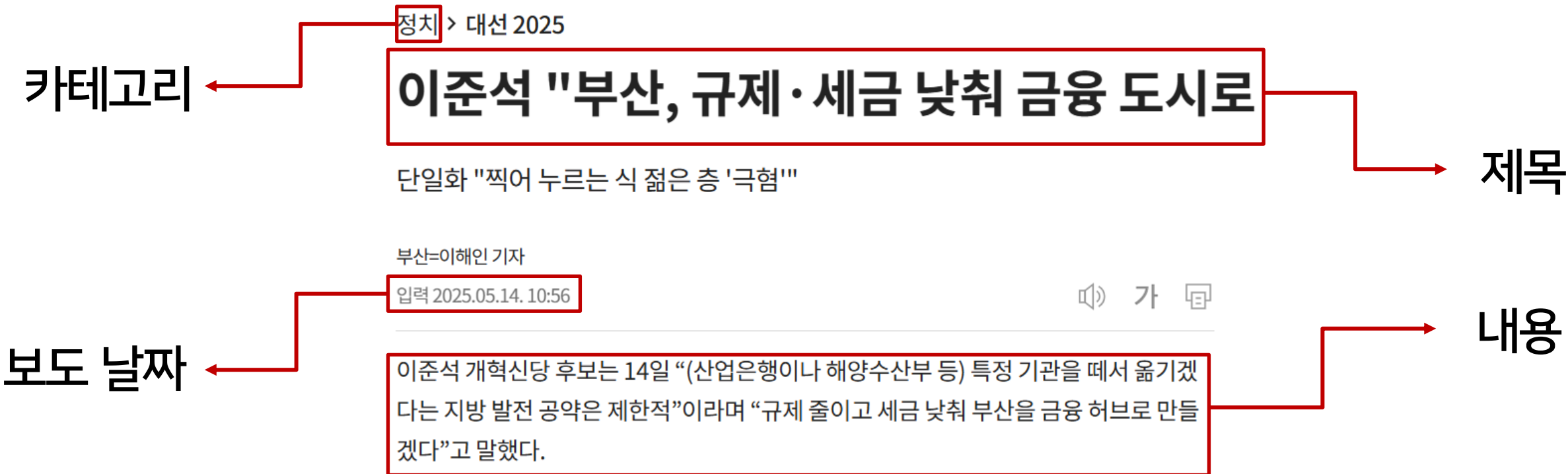
기간을 5주로 선택한 이유 : 미디어 아젠다와 대중 관심의 시간 의존성 정량화 - Sebastián Pinto 외 (2018) 참고



## 03 활용 데이터

### 뉴스 기사 데이터 | 크롤링

네이버 뉴스를 기사 수집 채널로 하여  
10개의 언론사에서 탄핵 선고일 전후 5주 간 보도한 기사를 크롤링



# 03 활용 데이터

## 뉴스 기사 데이터 | 데이터셋

크롤링을 통해 아래와 같은 형식의 **뉴스 기사 데이터셋**이 완성됨

	Title	Reporter	Date	Subject	Article
0	'G-패스' 어린이..	김호 기자	2025.03.21	뉴스9(광주)	광주시가 올해부터..
1	우상혁, 세계실내..	손기성 기자	2025.03.21	종합	한국 남자 높이뛰..
...	...	...	...	...	...
19032	5년간 성매매..	김보담 기자	2025.04.18	사회	성매매 업주들로..
19033	주말 또 전국 비..	김세현 기자	2025.04.18	뉴스 7	오늘도 고온 현상..

KBS 뉴스 기사 데이터셋

## 03 활용 데이터

### 국회 회의록 데이터 | 기간 선정

특정 시기나 주제에 한정된 발언만으로도 정치적 성향과 언어적 편향을 유의미하게 분석할 수 있음을 바탕으로, 본 연구는 **1년치 국회회의록**만으로도 충분한 분석 타당성을 갖춘다고 판단함

참고 문헌 : Gentzkow & Shapiro (2010), What Drives Media Slant? Evidence from U.S. Daily Newspapers  
박상현(2024), 주요 일간지의 20대 대선 보도에서 드러나는 정치적 편향성 분석



선행 연구들을 바탕으로

본 연구는 **제22대 국회**의 **1년치 회의록**만을 분석 대상으로 선정함

## 03 활용 데이터

### 국회 회의록 데이터 | API 데이터 수집

열린국회정보에서 제공하는 **국회 회의록 API**를 활용하여  
제22대 국회의 1년치 국회본회의, 상임위원회, 특별위원회 회의록을 다운로드

제415회-국토교통제3차(2024년6월25일) 1

제415회 국회 (임시회) 국토교통위원회회의록 제 3 호

국 회 사 무 처

일 시 2024년6월25일(화)

장 소 국토교통위원회회의실

#### 의사일정

- 전세사기피해자 지원 및 주거안정에 관한 특별법 일부개정법률안(김성원 의원 대표발의)(의안 번호 2200373)
- 전세사기피해자 지원 및 주거안정에 관한 특별법 일부개정법률안(염태영 의원 대표발의)(의안 번호 2200825)
- 전세사기피해자 지원 및 주거안정에 관한 특별법 일부개정법률안(허종식 의원 대표발의)(의안 번호 2200846)
- 전세사기 피해자 지원 대책 관련 청문회



PDF 형태로 존재하는 회의록에서  
**텍스트만 추출**하는 과정을 거침

### 03 활용 데이터

#### 국회 회의록 데이터 | 데이터셋

PDF 텍스트 추출을 통해 아래와 같은 형식의 **국회 회의록 데이터셋**이 완성됨

	대수	회의 날짜	회의종류명	회의명	텍스트
1	22	2024-06-25	상임위원회	제22대 제415회..	의사일정 1. 전세사기피해..
2	22	2024-06-25	상임위원회	제22대 제415회..	의사일정 1. 전세사기 등 주..
...	...	...	...	...	...
766	22	2025-04-02	국회본회의	제22대 제423회..	의사일정 1. 합성생물학육..

제22대 국회의 1년치 국회본회의, 상임위원회, 특별위원회 회의록

## 03 활용 데이터

### 국회의원 명단 데이터 | 회의록 한계점



국회 회의록은 정치적 성향을 명확하게 나타내는 자료이지만  
**발언자들의 소속 정당이 명확히 표시되지 않는 한계점** 존재함



국회의원 명단 데이터를 기반으로,  
발언자와 소속 정당 정보를 매칭하여 **각 발언자의 정치 성향을 분류**하고  
회의록 내 발언들을 **진보·보수 성향으로 나누는 기반**을 마련함

# 03 활용 데이터

## 국회의원 명단 데이터 | 데이터셋

대한민국국회 홈페이지에서 수집한 제 22대 국회의원 명단 데이터셋은 아래와 같음

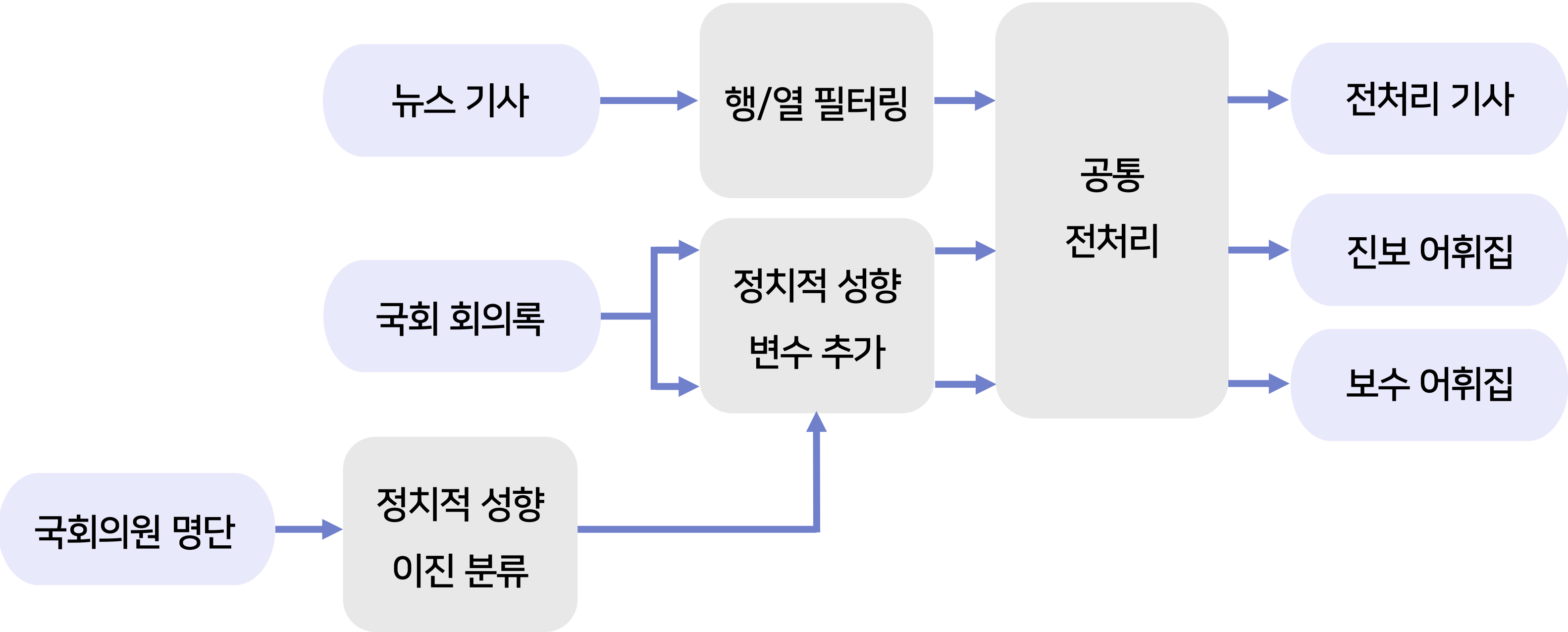
	구성	위원회	위원명	소속정당	...
1	위원장	국회운영위원회	박찬대	더불어민주당	...
2	간사	국회운영위원회	박성준	더불어민주당	...
...	...	...	...	...	...
484	위원	산불피해지원..	임종득	국민의힘	...

제22대 국회의 1년치 국회본회의, 상임위원회, 특별위원회 회의록 (국민의힘과 더불어민주당만 활용)



# 03 활용 데이터

## 데이터 전처리 | 전처리 흐름도



## 03 활용 데이터

### 개별 전처리 | 뉴스 기사 데이터

#### 카테고리 필터링 (행 필터링)

날씨, 스포츠, 연예 등 정치적 편향성 분석과  
**관계성이 낮은** 불필요한 카테고리의 샘플은 제거함

[조선일보]

유효한 카테고리: 조선경제, 정치, 국제, 사회, 부동산

제외한 카테고리: 문화·라이프, English, 영상·포토, 스포츠, 스페셜,  
국, 아무튼, 주말

#### 열 필터링

분석에 불필요한 제목, 보도 날짜, 기자, 카테고리를  
제거하고 **본문 내용**만 활용하기로 함

~~카테고리~~

정치 > 대선 2025

이준석 "부산, 규제·세금 낮춰 금융 도시로

단일화 "찍어 누르는 식 젊은 층 '극혐'"

제목

~~보도 날짜~~

부산=이해인 기자

입력 2025.05.14. 10:56

가

이준석 개혁신당 후보는 14일 "(산업은행이나 해양수산부 등) 특정 기관을 떼서 옮기겠  
다는 지방 발전 공약은 제한적"이라며 "규제 줄이고 세금 낮춰 부산을 금융 허브로 만들  
겠다"고 말했다.

내용

## 03 활용 데이터

### 개별 전처리 | 국회의원 명단 데이터 & 국회회의록 데이터

#### 국회의원 명단 데이터 - 정치적 성향 분류

의원별 소속 정당을 기준으로 하여 각 국회의원을 **진보/보수 진영으로 이진 분류하였고,**  
이는 국회 회의록 데이터의 **발언자의 정치 성향을 매핑하는 데 기반이 됨**

#### 국회회의록 데이터 - 정치적 성향 열 추가

진보·보수로 이분화한 국회의원 정치적 성향을 기준으로  
**각 의원이 발언한 내용에 정치 성향 정보를 매핑**하여 새로운 열을 추가함

## 03 활용 데이터

### 공통 전처리 | 토큰화 & 형태소 분석 및 어근 추출

#### 토큰화

바른 AI 형태소 분석기를 활용하여  
형태소 단위로 분리하고, 어근 중심의 토큰화 과정을 수행함



#### 형태소 분석 및 어근 추출

품사 태깅을 통해 각 형태소에 품사를 부착하고,  
분석 목적에 따라 명사·동사·형용사만 활용하도록 조사·접속사 등을 제거하여 필터링을 진행함

# 03 활용 데이터

## 공통 전처리 | TF-IDF 기반 불용어 제거

### 불용어 제거

전체 기사 및 국회 회의록 데이터에서 TF-IDF 기반  
영향력 지수 하위 5% 단어를 추출하여 불용어로 제거함

강남부  
강남분  
강남세무서  
강동군  
강동근  
강동을  
강동정치학원  
강령하  
강박하  
강변점

개과천선하  
개곶밭  
개념어  
개다리춤  
개두릅  
개떡  
개똥  
개리  
개발권  
개발전

결과안  
결과하  
결곶값  
결대로  
결선일  
결정난  
결정안  
결찰술  
결투하  
결피

추출된 불용어 예시

04

국회 회의록 편향성

유의성 검정

## 04 국회 회의록 편향성 유의성 검정

### 유의성 검정



토큰화 및 임베딩을 통해 생성한 국회 회의록 어휘집의 진보/보수 간 차이가 **통계적으로 유의미한지**

**MANOVA 검정**을 통해 확인하여 정치 편향 지수 산출의 타당성을 확보하고자 함

## 04 국회 회의록 편향성 유의성 검정

### 유의성 검정 | ① 빈도 기반 어휘집의 편향성

MANOVA(Multivariate Analysis of Variance)

여러 종속 변수에 대해 집단 간 차이를 동시에 검정하는 기법으로,  
정규성 및 등분산성 등의 가정을 충족해야 결과를 신뢰할 수 있음



빈도 기반 진보/보수 어휘집에 대한 MANOVA 검정하고자 하였으나  
다변량 정규성 검사 및 Box's M 검정 결과 정규성 및 등분산성을 만족하지 않았기 때문에  
비모수적 다변량 분석 방법인 Permutational MANOVA 검정을 사용하기로 함



## 04 국회 회의록 편향성 유의성 검정

### 유의성 검정 | ① 빈도 기반 어휘집의 편향성

#### PERMANOVA (Permutational MANOVA)

거리와 순열을 기반으로 집단 간 차이를 검정하는 비모수적 다변량 분산분석 기법으로,  
가정이 느슨하여 다변량 정규성이나 등공분산성을 만족하지 않아도 사용 가능함



#### Permutation Test 결과

$H_0$  : 진보/보수 성향 레이블과 단어 빈도는 무관하다.

$H_1$  : 진보/보수 성향 레이블과 단어 빈도는 관련이 있다.

p-value = 0.001 (< 0.05)

토큰화 된 의원들의 발언은 그들의 정치 성향과 통계적으로 유의하게 관련이 있다고 해석 가능

## 04 국회 회의록 편향성 유의성 검정

### 유의성 검정 | ② 의미 기반 어휘집의 편향성



임베딩 기반 국회 회의록의 진보/보수 데이터 또한  
정규성 및 등분산성을 만족하지 않아 **PERMANOVA 검정**을 진행함



#### Permutation Test 결과

$H_0$  : 진보/보수 어휘집 내 임베딩 벡터의 분포가 같다.

$H_1$  : 진보/보수 어휘집 내 임베딩 벡터의 분포에 유의한 차이가 있다.

p-value = **0.001** (< 0.05)

임베딩 된 의원들의 발언은 그들의 **정치 성향과 통계적으로 유의하게 관련이 있다**고 해석 가능

## 04 국회 회의록 편향성 유의성 검정

### 유의성 검정 | ② 의미 기반 어휘집의 편향성



임베딩 기반 국회 회의록의 진보/보수 데이터 또한  
정규성 및 등분산성을 만족하지  PERMANOVA 검정을 진행함

즉, 국회 회의록 진보 어휘집과 보수 어휘집 간에 **유의미한 차이**를 보였으며  
국회 회의록 진보/보수 어휘집을 활용한 **정치 편향 지수 산출의 타당성을 확보함**

$H_0$  : 진보/보수 어휘집 내 임베딩 벡터의 분포가 같다.

$H_1$  : 진보/보수 어휘집 내 임베딩 벡터의 분포에 유의한 차이가 있다.

p-value = 0.001 (< 0.05)

임베딩 된 의원들의 발언은 그들의 **정치 성향과 통계적으로 유의하게 관련이 있다**고 해석 가능

05

# 정치 편향 지수 산출

## 05 정치 편향 지수 산출

### 정치 편향 지수 산출

지금부터 두 가지 방법으로  
정치 편향 지수를 산출해보자!



빈도 기반 유사도

단어의 출현 빈도를 측정하여  
각 어휘집과의 유사도를 측정



의미 기반 유사도

단어의 의미를 반영하여  
문맥적 유사도를 측정

## 05 정치 편향 지수 산출



### 정치 편향 지수 산출

#### 진보와 보수

두 어휘집으로 측정한 유사도?  
지금부터 두 가지 방법으로

정치 편향 지수를 산출해보자!

한 기사에 대하여 진보 어휘집과 보수 어휘집,

두 어휘집으로 유사도를 측정하면 두 가지 유사도가 생기며

두 가지 유사도의 차이를 통해 하나의 유사도를 생성할 수 있음

단어의 출현 빈도를 측정하여

각 어휘집과의 유사도를 측정

단어의 의미를 반영하여

문맥적 유사도를 측정

이를 바탕으로 유사도가 1에 가까우면 진보,

-1에 가까우면 보수가 되도록 설정함

## 05 정치 편향 지수 산출

### 빈도 기반 정치 편향 지수



토큰화된 국회 회의록의 **진보/보수 발언 어휘집**을 참조 코퍼스로 설정하고  
분석 대상인 기사 데이터의 **어휘 출현 빈도**와 비교하여 지수 산출

코퍼스(Corpus) : 말뭉치



TF - IDF

각 용어의 **문서별 가중치**를 계산하고  
가중치 곱 연산으로 편향 점수 산출



Log-odds Ratio

각 단어의 **상대적 사용 빈도**를 계산하고  
로그 오즈 비율로 변환하여 편향 점수 산출

# 05 정치 편향 지수 산출

## 빈도 기반 정치 편향 지수 | ① TF - IDF

### TF - IDF

문서 내 단어의 등장 빈도와 전체 문서에서의 희귀성을 고려하여  
해당 단어의 중요도를 수치화 하는 기법

#### ① 정당별 TF-IDF 계산

진보/보수 어휘집을 활용하여 두 정당의 TF-IDF 계산

#### ② 단어별 편향 가중치 생성

(진보 TF-IDF) - (보수 TF-IDF)를 계산하여 가중치 생성

#### ③ 정치 편향 지수 생성

기사별 단어 등장 빈도와 TF-IDF 가중치 결합



# 05 정치 편향 지수 산출

## 빈도 기반 정치 편향 지수 | ① TF - IDF

TF-IDF 기반 정치 편향 지수 계산 결과,  
아래와 같은 형식의 언론사별 데이터셋이 완성됨

	원본 기사	편향 지수
1	대통령 경호처 직원, 동료와 다투다 경찰관 폭행에 ...	0.240929
2	HD현대, 정주영 창업자 24주기 추모행사 ...	0.224596
3	中, '마약사범' 캐나다 4명 사형 집행 ...	0.178544
...	...	...

TV조선 기사 데이터에 대한 TF-IDF 정치 편향 지수 데이터셋

## 05 정치 편향 지수 산출

### 빈도 기반 정치 편향 지수 | ② Log-odds Ratio

#### Log-odds Ratio

특정 단어가 두 집단에서 사용되는 **상대적 비율**을 비교하여  
그 비율의 로그값을 통해 **단어의 집단 간 편향 정도**를 수치화하는 기법

#### ① 단어별 확률 계산

진보/보수 어휘집에서의 상대적 빈도를 기반으로 확률을 계산

#### ② 단어별 Log-odds 계산

진보 단어면 log-odds 기여도를 더하고, 보수 단어면 빼는 방식

#### ③ 정치 편향 지수 생성

단어 빈도와 log-odds를 결합하고 문서의 토큰 수로 나누는 방식

## 05 정치 편향 지수 산출

### 빈도 기반 정치 편향 지수 | ② Log-odds Ratio

#### Log-odds Ratio

특정 단어가 두 집단에서 사용되는 **상대적 비율**을 비교하여  
그 비율의 로그값을 통해 **단어의 집단 간 편향 정도**를 수치화하는 기법

① 단어별 확률 계산



진보/보수 이념에서의 상대적 빈도를 기반으로 확률을 계산

이때, **카이제곱 검정**을 활용하여 진보와 보수 간에 단어 사용 빈도 차이가

**통계적으로 유의미한 단어만** 활용하여 **진보 토큰 Pool**과 **보수 토큰 Pool**을 생성하였고

이를 바탕으로 Log-odds Ratio 기반 빈도 유사도를 계산함

③ 정치 편향 지수 생성

단어 빈도와 log-odds를 결합하고 문서의 토큰 수로 나누는 방식

# 05 정치 편향 지수 산출

## 빈도 기반 정치 편향 지수 | ② Log-odds Ratio

Log-odds Ratio 기반 정치 편향 지수 계산 결과,  
아래와 같은 형식의 언론사별 데이터셋이 완성됨

	제목	원본 기사	언론사	편향지수
1	정부...	미국 에너지부가 민감국가 목록에 한국을 포함...	JTBC	4.501283
2	[르포] ...	[앵커] 포천 오폭 사고로 인근 가축 농가도 직격탄...	JTBC	2.935053
3	헌재 ...	■ 방송 : JTBC 논/쟁 / 진행 : 오대영■...	JTBC	4.706960
...	...	...	...	...

Log-odds Ratio 정치 편향 지수 데이터셋

## 05 정치 편향 지수 산출

### 빈도 기반 정치 편향 지수 | 정규화

산출된 정치 편향 지수에 대하여  
두 가지 변환을 통해 정규화를 진행함

#### Quantile Transformation

데이터의 분포를 변환하여  
이상치의 영향을 줄이는 전처리 기법

#### -1~1 선형 변환

모든 값을 -1과 1 사이로  
선형적으로 변환하는 스케일링 기법

# 05 정치 편향 지수 산출

## 빈도 기반 정치 편향 지수 | 완성된 데이터셋

TF-IDF 정치 편향 지수와 Log-odds Ratio 정치 편향 지수에 정규화를 수행하여  
아래와 같은 형식의 데이터셋을 완성함

	제목	TF-IDF 편향 지수	Log-odds Ratio 편향 지수	언론사
1	"미국인의 英시민권 신청 급증... 트럼프 효과"...	-0.3892018	-0.041884	TV조선
2	'尹 궤변' 깨질 때마다 탄성 .. "이렇게 당연한 걸"...	0.2426419	0.1889489	MBC
3	'尹 심판' 20석 일반인 방청 신청에 9만 6000명...	-0.4621811	-0.172013	TV조선
...	...	...	...	...

# 05 정치 편향 지수 산출

## 의미 기반 정치 편향 지수



국회 회의록의 **진보/보수 발언 어휘집**을 참조 코퍼스로 설정하고  
분석 대상인 기사 데이터의 **의미 표현**과 비교

코퍼스(Corpus) : 말뭉치



KoSBERT

Korean Sentence Bert



KoSimCSE

Korean Simple Contrastive  
Learning of Sentence Embeddings

## 05 정치 편향 지수 산출

### 의미 기반 정치 편향 지수 | 유사도 측정

각 데이터셋의 임베딩 유사도 측정

국회 회의록 임베딩과 진보/보수 임베딩의 각 코사인 유사도를 측정하고  
각 유사도의 차이를 통해 하나의 편향 지수를 생성

$$\left( \text{진보 임베딩 유사도} \right) - \left( \text{보수 임베딩 유사도} \right) \\ = \left( \text{최종 편향 지수} \right)$$



# 05 정치 편향 지수 산출

## 의미 기반 정치 편향 지수 | 지수 산출 결과

최종적으로 KoSBERT와 KoSimCSE 기반 편향 지수가 아래와 같이 측정됨

	제목	편향 지수
1	정부 "한미 양국, '민감국가' ...	0.00418
2	[르포] "유산하고 죽고.. ...	0.00029
3	'현재 선고 늦어지는 이유는? ...	0.00498
...	...	...

KoSBERT 편향 지수 측정 결과

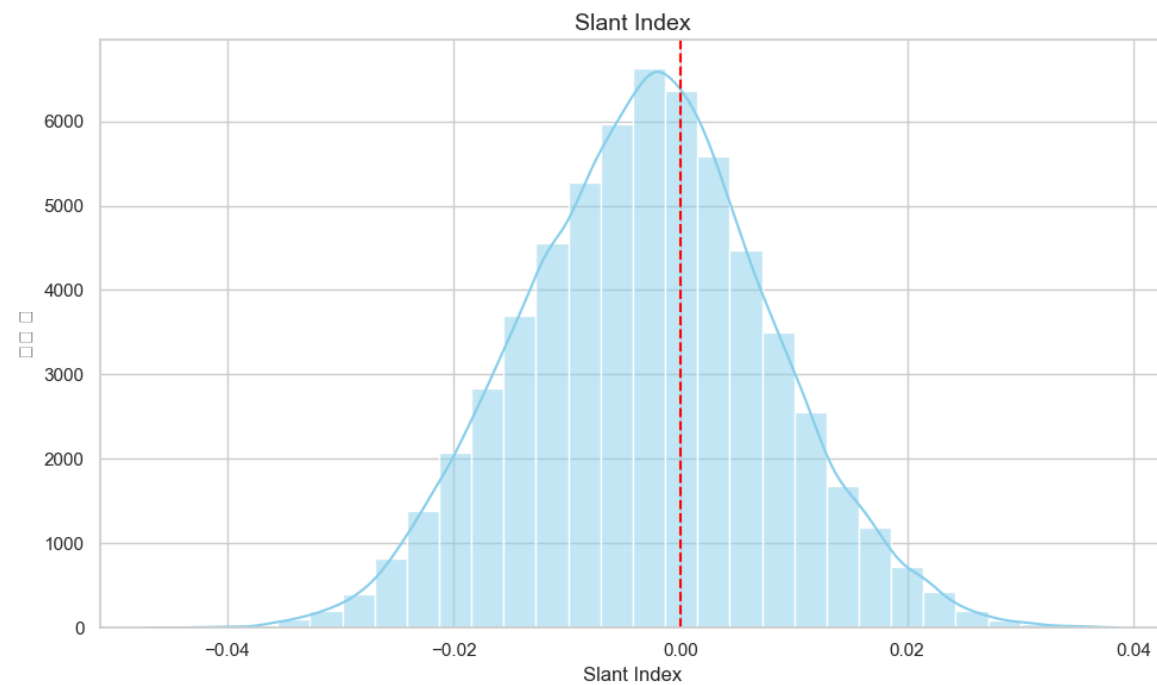
	제목	편향 지수
1	정부 "한미 양국, '민감국가' ...	0.00532
2	[르포] "유산하고 죽고.. ...	-0.00735
3	'현재 선고 늦어지는 이유는? ...	0.02190
...	...	...

KoSimCSE 편향 지수 측정 결과

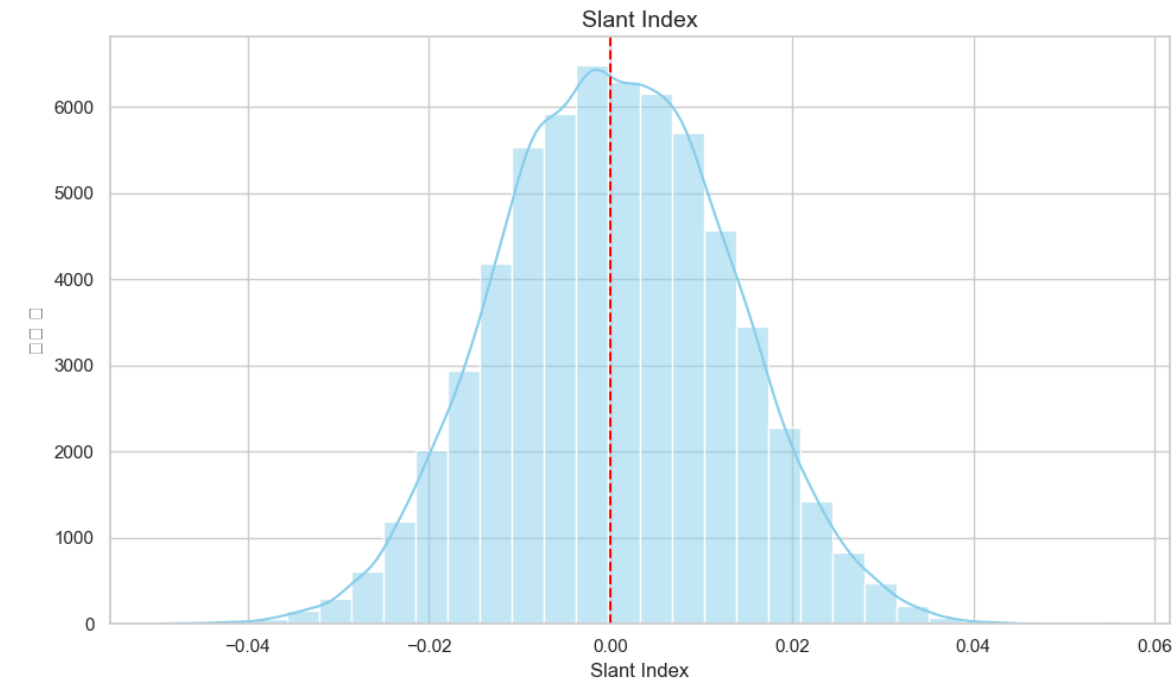
## 05 정치 편향 지수 산출

### 의미 기반 정치 편향 지수 | 지수 산출 결과

의미 기반 편향 지수를 각각 시각화한 결과, 아래와 같이 정규 분포와 유사한 형태를 보임



KoSBERT 기반 편향 지수 분포



KoSimCSE 기반 편향 지수 분포

대부분의 값이 **-0.04에서 0.06 사이에 밀집**되어 있기 때문에  
빈도 기반 정치 편향 지수와 동일하게 **Quantile Transformer**를 통한 **정규화**로 문제를 해결

## 05 정치 편향 지수 산출



의미 기반 정치 편향 지수 | 지수 산출 결과

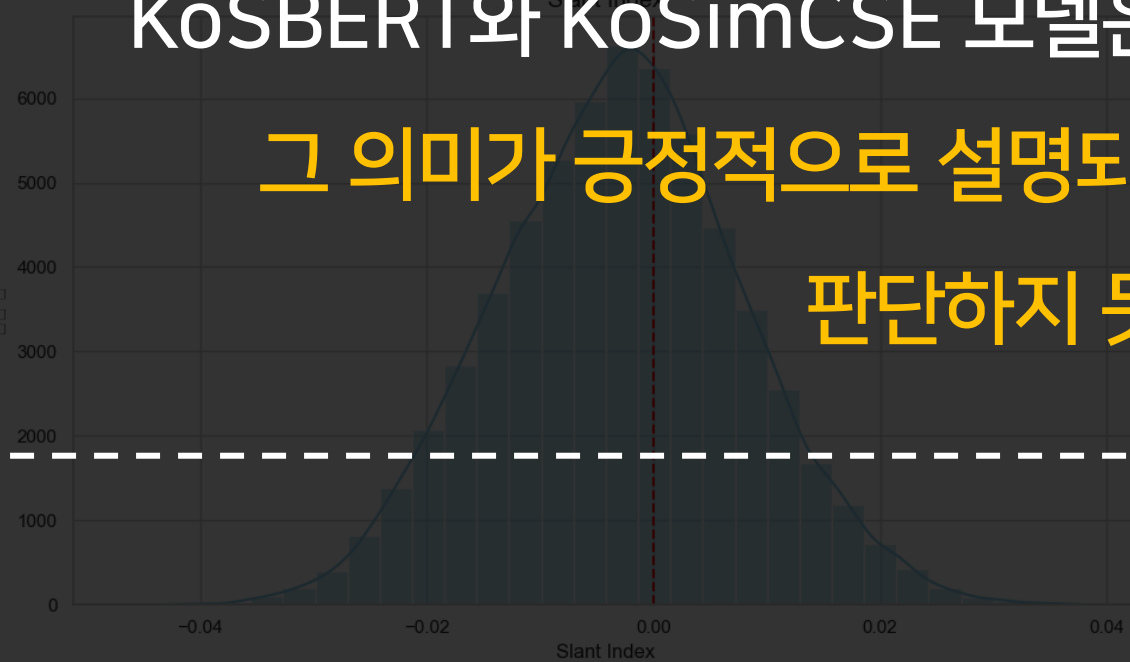
### 의미 기반 모델의 한계점

의미 기반 편향 지수를 각각 시각화한 결과, 아래와 같이 정규 분포와 유사한 형태를 보임

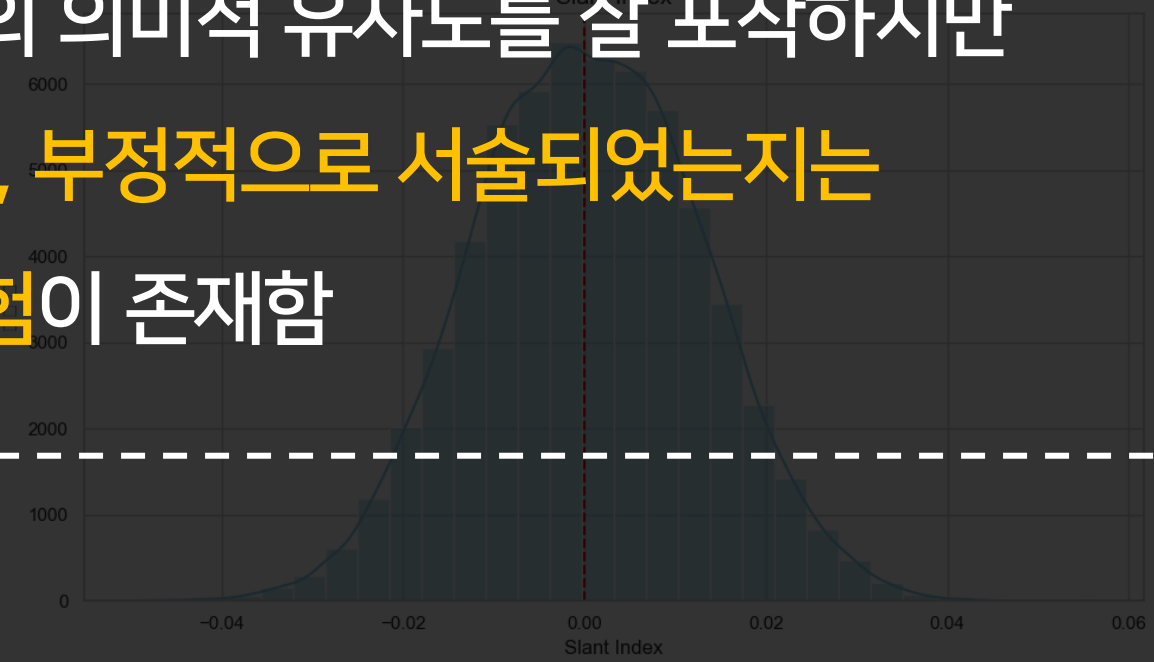
KoSBERT와 KoSimCSE 모델은 문장의 의미적 유사도를 잘 포착하지만

그 의미가 긍정적으로 설명되었는지, 부정적으로 서술되었는지는

판단하지 못할 위험이 존재함



KoSBERT 기반 편향 지수 분포



KoSimCSE 기반 편향 지수 분포

긍정적으로 기술되었는지, 부정적으로 비판했는지를 파악하기 위해

대부분의 값이 -0.04에서 0.06 사이에 밀집되어 있기 때문에

감성 분석을 추가로 도입

빈도 기반 정치 편향 지수와 동일하게 GPT-4를 통한 정규화로 문제를 해결

## 05 정치 편향 지수 산출

### ■ 감성 분석

KcELECTRA-base-v2022

한국어에 특화된 사전학습 모델인 KcELECTRA를 기반으로 한 감성 분류 모델로  
긍정 / 부정 / 중립 감정을 포함하고 있는지를 판단

감성 분석 결과를 긍정은 1, 부정은 -1, 중립은 0으로 수치화한 후  
이를 기존의 정치 편향 지수와 곱하여 최종 성향 지수로 반영

진보 × 긍정 = 진보    진보 × 부정 = 보수  
보수 × 긍정 = 보수    보수 × 부정 = 진보

# 05 정치 편향 지수 산출

## 감성 분석

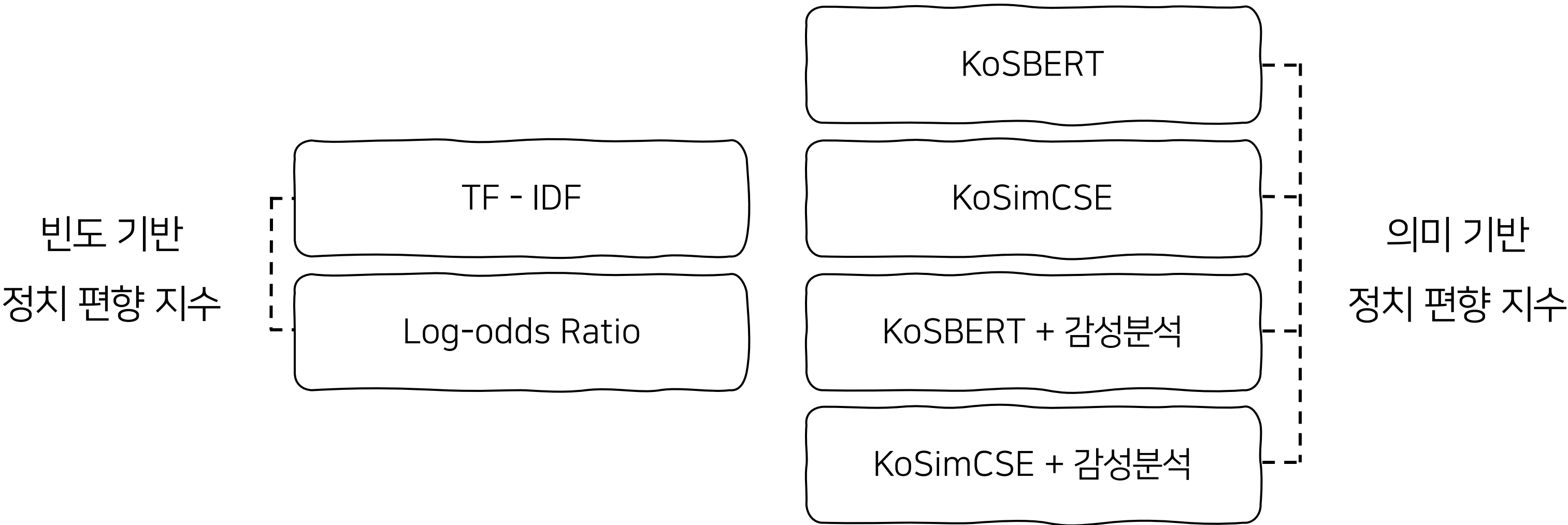
최종적으로 의미 기반 편향 지수와 감성 분석 결과를 결합하여 아래와 같이 편향 지수가 측정됨

	제목	KoSBERT 편향 지수	KoSimCSE 편향 지수	감성 분석 결과	KoSBERT + 감성 분석	KoSimCSE + 감성 분석
1	정부 "한미 양국, '민감국가' ...	0.0042	0.00532	-1	-0.0042	-0.00532
2	[르포] "유산하고 죽고.. ...	0.0003	-0.00735	1	0.0003	-0.00735
3	'현재 선고 늦어지는 이유는? ...	0.0050	0.02190	1	0.0050	0.02190
...	...	...	...	...	...	...

# 05 정치 편향 지수 산출

## 최종 지수 산출

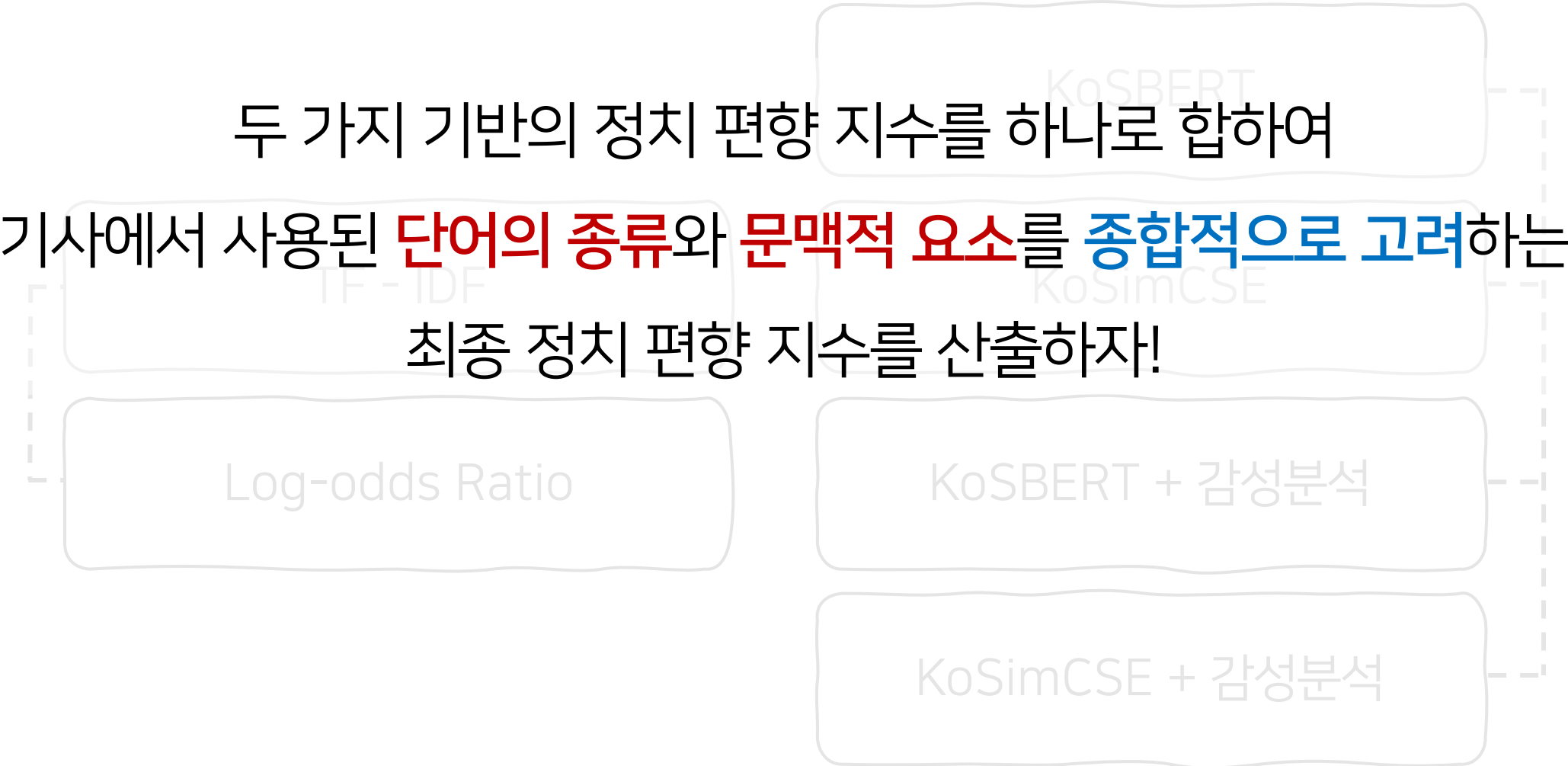
측정된 정치 편향 지수의 종류는 아래와 같이 구성됨



# 05 정치 편향 지수 산출

## 최종 지수 산출

측정된 정치 편향 지수  는 아래와 같이 구성됨



## 05 정치 편향 지수 산출

### 최종 지수 산출



측정된 정치 편향 지수의 종류는 아래와 같이 구성됨

이때, **가중평균**을 통해 하나의 정치 편향 지수를 생성

$$\alpha \times \left( \begin{array}{c} \text{빈도 기반} \\ \text{정치 편향 지수} \end{array} \right) + (1 - \alpha) \times \left( \begin{array}{c} \text{의미 기반} \\ \text{정치 편향 지수} \end{array} \right)$$

빈도 기반  
정치 편향 지수

KoSBERT

KoSimCSE

의미 기반  
정치 편향 지수

이후, **새로운 데이터셋**에 대하여 여덟 가지 조합의

정치 편향 지수를 산출하는 과정을 통해 **가장 성능이 좋은 조합**과 **가중치**를 탐색

KoSimCSE + 감성분석



## 05 최종 정치 편향 지수 산출

### 실험용 데이터셋

국회회의록 데이터만 활용하여 발생하는 **과적합을 방지하고 일반화 성능을 평가**하기 위해  
다른 도메인이면서 정치 성향이 극명하게 나타나는  
인터넷 커뮤니티의 게시글을 **실험용 데이터**로 활용함

크롤링을 통해 진보와 보수의 비율이 50:50인 **실험용 데이터셋**<sup>👉</sup>이 완성됨

	성향	제목	본문
1	보수	사전투표자 총...	이런데도 사전투표가 조작됐다는 교회 사기꾼...
...	...	...	...
100	보수	현재 이준석과...	"신용도" 총리자리든, 당대표자리든 머든 약속한...

# 05 최종 정치 편향 지수 산출

## 실험 결과 | 가장 성능이 좋은 조합



실험용 데이터에 대한 **여덟 가지 조합**의 **정치 편향 지수**를 측정하고 정확도를 검사한 결과  
최종적으로 **Log-odds Ratio**와 **KoSBERT**의 조합이 최고의 조합으로 선정됨

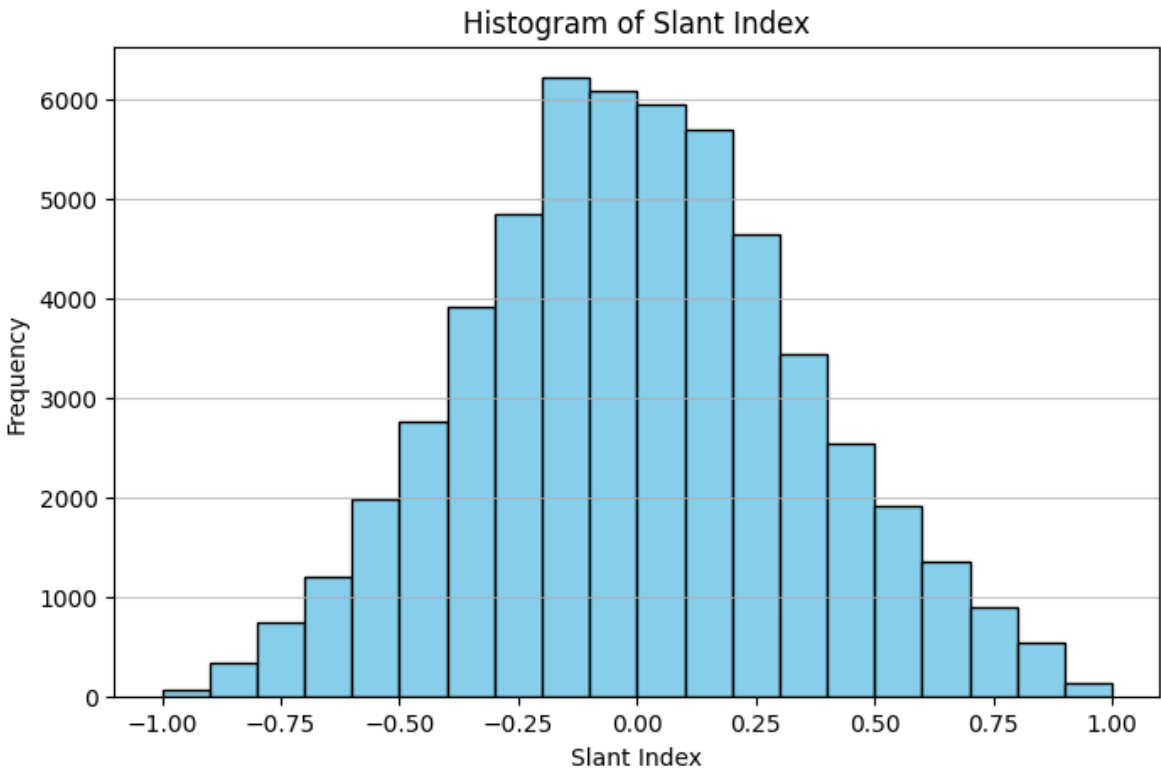
$0.577 \times (\text{Log-odds Ratio}) + 0.423 \times (\text{KoSBERT})$

빈도 기반 정치 편향 지수	의미 기반 정치 편향 지수	$\alpha$	MSE	분류 정확도
Log-odds Ratio	KoSBERT	0.577	0.807	70%

# 05 최종 정치 편향 지수 산출

## 최종 편향 지수 산출 결과

	제목	원본 기사	최종 편향 지수
1	윤석열 정부 1,060일...3년 못 채우고 단명 자초	윤석열 전 대통령은 2년 11개월 가까운 임기 동안 부인 김건희 여사 관련 의혹...	0.844
...	...	...	...
75	경찰, 광화문 서십자각 농성장서 전농 트랙터 견인...	현재 비상행동 등 탄핵찬성 단체들이 도로를 점거하고 경찰과 대치하는 등 ...	- 0.419
...	...	...	...



앞서 구한 산출식을 바탕으로 각 뉴스 기사 별로 **최종 편향 지수**를 계산한 결과  
위와 같이 **정규분포와 유사한 형태로** 나타남

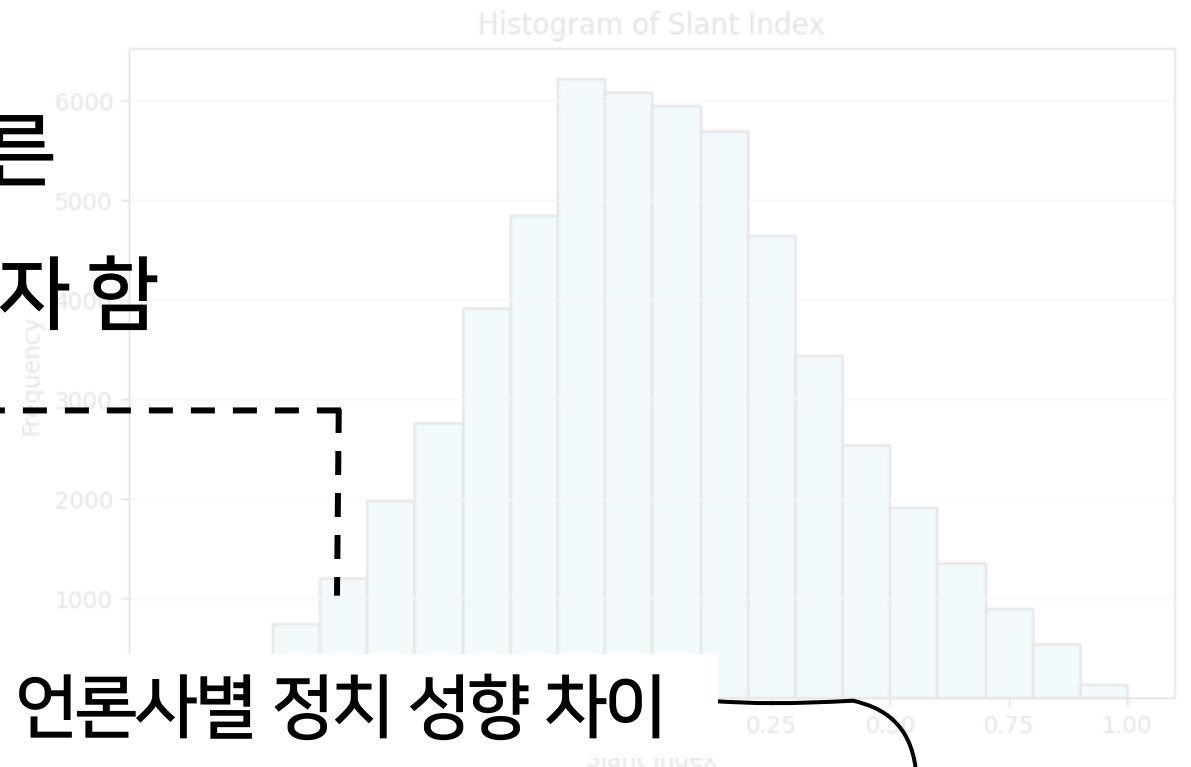
# 05 최종 정치 편향 지수 산출

## 최종 편향 지수 산출 결과



정치 편향 지수 산출에 따른  
의문사항 두 가지를 검정하고자 함

	제목	원본 기사	최종 편향 지수
1	윤석열 정부 1,060일...3년 못 채우고 단명 자초	윤석열 전 대통령은 2년 11개월 가까운 임기 동안 부인 김건희 여사 관련 의혹	0.844
...	...	...	...
75	경찰, 광화문 서십자각 농성 장서 전농 트랙터 견인	현재 비상행동 등 탄핵찬성 단체들이 도로를 점거하고 경찰	- 0.419
...	...	...	...



👉 중립 기사 범위  
몇 점의 기사를 중립적이라고  
판단할 수 있을까?

👉 언론사별 정치 성향 차이  
언론사별로 평균적으로  
어떤 정치적 성향을 보일까?

## 05 최종 정치 편향 지수 산출

### 최종 편향 지수 산출 결과 | 중립 기사 범위

기사의 정치 편향 지수가 중립인지 아닌지를 판단하기 위해  
전체 데이터의 평균과 표준편차를 바탕으로 95% 신뢰구간을 계산함



95% 신뢰구간 도출 결과

$$(\text{표본 평균}) \pm 1.96 \times (\text{표준 오차}) = [-0.003, 0.0028]$$

**[-0.003, 0.0028]** 범위 안의 편향 지수를 갖는 기사는

**통계적으로 중립**이라고 판단 가능

## 05 최종 정치 편향 지수 산출

### 최종 편향 지수 산출 결과 | 언론사별 정치 성향 차이



#### 언론사별 정치 편향 지수 검정 결과 판단 방법

- ① One-sample t-test : 평균이 0과 통계적으로 다른지를 평가
- ② 평균 편향 지수 : 평균적으로 진보적인 성향인지 보수적인 성향인지를 평가
- ③ 95% Bootstrap 신뢰구간 : 평균 편향 지수의 통계적 유의성 평가를 위한 구간

T-test p-value가 0.05보다 작고 신뢰구간이 0을 포함하지 않으면 **중립이 아니고**

T-test p-value가 0.05보다 크거나 같고 신뢰구간이 0을 포함하면 **중립**이라고 판단 가능하며,

중립이 아닐 때 평균이 0보다 크면 **진보 성향**, 작으면 **보수 성향**이라고 판단 가능

# 05 최종 정치 편향 지수 산출

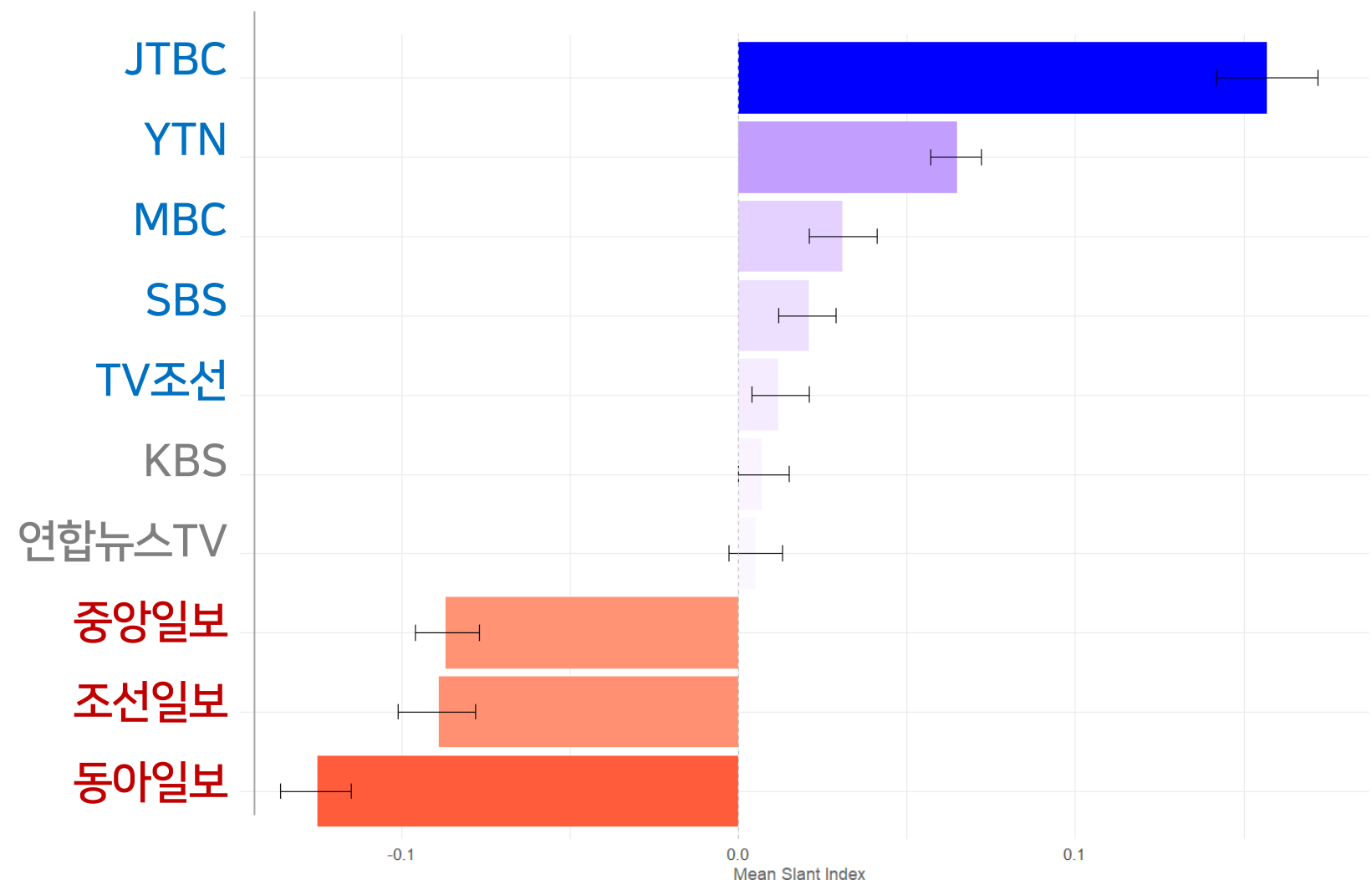
## 최종 편향 지수 산출 결과 | 언론사별 정치 성향 차이

언론사별 평균 정치 편향 지수 검정 결과

	언론사	One-sample t-test (p-value)	Mean	95% CI	진보/보수
1	JTBC	0.000	0.157	[0.142, 0.172]	진보
2	YTN	0.000	0.065	[0.057, 0.072]	진보
3	MBC	0.000	0.031	[0.021, 0.041]	진보
4	SBS	0.000	0.021	[0.012, 0.029]	진보
5	TV조선	0.006	0.012	[0.004, 0.021]	진보
6	KBS	0.050	0.007	[0.000, 0.015]	중립
7	연합뉴스TV	0.233	0.005	[-0.003, 0.013]	중립
8	중앙일보	0.000	-0.087	[-0.096, -0.077]	보수
9	조선일보	0.000	-0.089	[-0.101, -0.078]	보수
10	동아일보	0.000	-0.125	[-0.136, -0.115]	보수

# 05 최종 정치 편향 지수 산출

## 최종 편향 지수 산출 결과 | 언론사별 정치 성향 차이



앞서 구한 언론사별 정치적 성향 지수의 평균 및 신뢰구간을 시각화 한 결과,  
통상적으로 진보/보수라고 대표되는 언론사들의 성향을 잘 담아낸 것을 확인할 수 있음

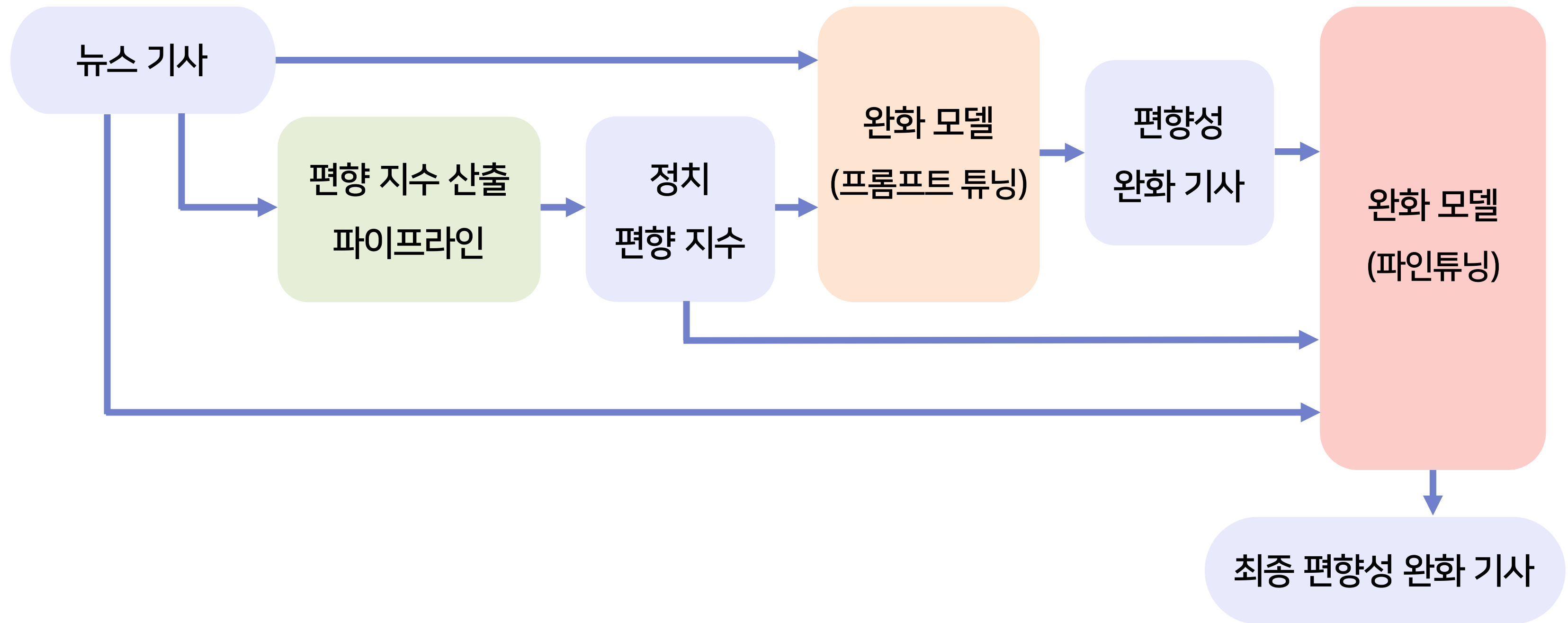


06

# 편향성 완화 모델

## 06 편향성 완화 모델

### 편향성 완화 모델 튜닝 흐름도



## 06 편향성 완화 모델

### 프롬프트 튜닝

#### 프롬프트 튜닝 (Prompt Tuning)

대규모 언어 모델이 특정 작업을 더 잘 수행하도록  
**프롬프트(입력 텍스트)**를 정교하게 구성하는 기술

모델 : GPT-4o

OpenAI에서 개발한  
**GPT-4** 계열의  
텍스트 기반 대규모 언어 모델

기법 : Few-shot learning

**예시(few examples)**를  
프롬프트에 작성하여  
모델의 성능을 향상시키는 학습 기법

GPT-4o 웹 인터페이스를 활용하였으며,

더 높은 **정밀도와 일관성**을 확보할 수 있도록 **Few-shot learning** 기법을 사용하기로 함

## 06 편향성 완화 모델

### 프롬프트 튜닝 | GPT-4o

#### GPT-4o 프롬프트 튜닝

----- 예시 -----

{example\_input}

----- 변환 -----

다음 기사를 정치적으로 중립적인 방향으로 완화해주세요.

완화된 기사는 원문 기사 분량의 최소 60% 이상이 되도록 해주세요.

즉, 너무 짧게 요약하지 말고 전체적인 길이와 정보량을 최대한 유지해주세요.

## 06 편향성 완화 모델

### 프롬프트 튜닝 | GPT-4o

#### GPT-4o 프롬프트 튜닝

----- 예시 -----

{example\_input}

----- 변환 -----

다음 기사를 정치적으로 중립적인 방향으로 주세요.

완화된 기사는 원문 **편향 지수가 0점에 가까운 중립 기사 5개**를 제공하여

즉, 너무 짧게 요약하지 말고 **Few-shot learning 기법**을 활용함

## 06 편향성 완화 모델

### 프롬프트 튜닝 | GPT-4o

#### GPT-4o 프롬프트 튜닝



----- 예시 -----  
{example input}  
----- 변환 -----

시행착오를 통해 적절한 프롬프트를 설계하였으며  
글이 과도하게 생략되는 것을 방지하기 위해 **구체적인 분량을 제시**하여 **정보량을 유지**하도록 함

다음 기사를 정치적으로 중립적인 방향으로 완화해주세요.  
완화된 기사는 **원문 기사 분량의 최소 60% 이상**이 되도록 해주세요.  
즉, 너무 짧게 요약하지 말고 **전체적인 길이와 정보량을 최대한 유지**해주세요.

## 06 편향성 완화 모델

### 프롬프트 튜닝 | GPT-4o

#### GPT-4o 프롬프트 튜닝 결과

	제목	원본 기사	편향 지수	완화 기사	완화 기사 편향 지수
1	'윤 대통령 명예훼손' 언론인 재판 시작... "공소 기각돼야"	오늘(31일) 서울중앙지방법원에서 열 린 재판에서 검찰은 부패·경제 범죄에 대해서만 수사를 개시할 수 있다며 이 사건은 기소 자체가 위법해 공소가 기 각돼야 한다고 ...	0.983	대통령 선거 기간 중 발생한 언론 보도 와 관련해 명예훼손 혐의로 기소된 언 론인들과 정치권 인사에 대한 첫 재판 이 서울중앙지방법원에서 열렸다. 피 고인 측은 공소 제기 절차의 ...	0.331
...	...	...	...	...	...
40	검찰, '사드 기밀 유출 혐의' 정의용 불구속 기소	문재인 정부 안보 고위직 인사 3명이 고고도미사일방어체계, 사드 배치 고 의 지연 의혹과 관련해 재판 ...	- 0.954	검찰은 고고도미사일방어체계(사드) 관련 정책 결정 과정에서 직권남용 및 비밀누설 등의 혐의를 받은 ...	- 0.153
...	...	...	...	...	...

GPT-4o 프롬프트 튜닝으로 모델이 기사를 **잘 완화시킴**을 확인함 🤗

## 06 편향성 완화 모델

### 프롬프트 튜닝 | GPT-4o

#### GPT-4o 프롬프트 튜닝 결과

	제목	원본 기사	편향 지수	완화 기사	완화 기사 편향 지수
1	'윤 대통령 명예훼손' 언론인 재판 시작... "공소 기각돼야"	오늘(31일) 서울중앙지방법원에서 열 린 재판에서 검찰은 부패·경제 범죄에 대해서만 수사를 개시할 수 있다며 이 사건은 기소 자체가 위법해 공소가 기 각돼야 한다고 ...	0.983	대통령 선거 기간 중 발생한 언론 보도 와 관련해 명예훼손 혐의로 기소된 언 론인들과 정치권 인사에 대한 첫 재판 이 서울중앙지방법원에서 열렸다. 피 고인 측은 공소 제기 절차의 ...	0.331
...	...	...	...	...	...
40	검찰, '사드 기밀 유출 혐의' 정의용 불구속 기소	문재인 정부 안보 고위직 인사 3명이 고고도미사일방어체계, 사드 배치 고 의 지연 의혹과 관련해 재판 ...	- 0.954	검찰은 고고도미사일방어체계(사드) 관련 정책 결정 과정에서 직권남용 및 비밀누설 등의 혐의를 받은 ...	- 0.153

점수 및 기사 내용을 종합적으로 고려하여 **완화가 잘 된 40개의 기사**를 추출하고

이를 이용하여 **GPT-3.5 Turbo** 모델을 **파인튜닝** 하기로 함



# 06 편향성 완화 모델

## 파인튜닝 | GPT-3.5 Turbo



### GPT-3.5 Turbo 파인튜닝 프로세스

01

데이터 준비

JSONL 파일 생성

02

모델 파인튜닝

파인튜닝 환경 구성 및 실행

03

완화 작업 수행

파인튜닝된 모델로  
기사의 편향성 완화 작업 수행

# 06 편향성 완화 모델

## 파인튜닝 | GPT-3.5 Turbo



데이터 준비

프롬프트 튜닝을 통해 마련한 40개의 완화 기사 데이터로

(input, output) 쌍으로 이루어진 JSONL 파일을 생성하여 학습 데이터 마련

Input Samples

Output Samples

	제목	원본 기사	편향 지수	완화 기사	완화 기사 편향 지수
1	'윤 대통령 명예훼손' 언론인 재판 시작... "공소 기각돼야"	오늘(31일) 서울중앙지방법원에서 열린 재판에서 검찰은 부패·경제 범죄에 대해서만 수사를 개시할 수 있다며 이 사건은 기소 자체가 위법해 공소가 기각돼야 한다고 ...	0.983	대통령 선거 기간 중 발생한 언론 보도와 관련해 명예훼손 혐의로 기소된 언론인들과 정치권 인사에 대한 첫 재판이 서울중앙지방법원에서 열렸다. 피고인 측은 공소 제기 절차의 ...	0.531
...	...	...	...	...	...

## 06 편향성 완화 모델

### 파인튜닝 | GPT-3.5 Turbo



#### 데이터 준비

##### Fine-tuning용 JSONL 파일

```
{ "messages": [  
  {"role": "system", "content": " 당신은 정치적으로 중립적인 뉴스로 문장을 변환하는 모델입니다. " },  
  { " role " : " user " , " content " : " 다음 기사를 정치적으로 중립적인 방향으로 완화해주세요.  
[제목] : '윤 대통령 명예훼손' 언론인들 재판 시작... "공소 기각돼야"  
[기사 원문] : 이 사건은 기소 자체가 위법해 공소가 기각돼야 한다고 주장 ... (중략) ...  
[편향 지수] : 0.97 (진보)"},  
  {"role": "assistant", "content": "피고인 측은 공소 제기 절차의 적법성에 의문을 제기 ... (중략) ..."} ] } ...
```

## 06 편향성 완화 모델

### 파인튜닝 | GPT-3.5 Turbo



데이터 준비

Fine-tuning용 JSONL 파일

```
{ "messages": [
```

```
  {"role": "system", "content": " 당신은 정치적으로 중립적인 뉴스로 문장을 변환하는 모델입니다. " }
```

```
  { "role": " user ", " content " : " 다음 기사를 정치적으로 중립적인 방향으로 완화해주세요.
```

```
[제목] : '윤 대통령 명예훼손' 언론인들 재판 시작 "공소 기간돼야"
```

```
[기사 원문] : 이 사건은 기소 자체가 위법해 공소가 기각돼야 한다고 주장 ... (중략) ...
```

```
[편향 지수] : 0.97 (진보)"},
```

```
  {"role": "assistant", "content": "피고인 측은 공소 제기 절차의 적법성에 의문을 제기 ... (중략) ..." } ] } ...
```

**system**은 모델의 역할이나 성격을 설정하는 프롬프트로  
중립적 뉴스 변환 모델임을 명시

## 06 편향성 완화 모델

### 파인튜닝 | GPT-3.5 Turbo



데이터 준비

Fine-tuning용 JSONL 파일

`{ "messages": [`  
"user"는 input 문장을 설정하는 프롬프트로 (제목, 기사 원문, 편향 지수)를 제공  
`{ "role": "system", "content": "당신은 정치적으로 중립적인 뉴스로 문장을 변환하는 모델입니다." },`

```
{ "role " : " user " , " content " : " 다음 기사를 정치적으로 중립적인 방향으로 완화해주세요.  
[제목] : '윤 대통령 명예훼손' 언론인들 재판 시작... "공소 기각돼야"  
[기사 원문] : 이 사건은 기소 자체가 위법해 공소가 기각돼야 한다고 주장 ... (중략) ...  
[편향 지수] : 0.97 (진보)" },
```

```
{ "role": "assistant", "content": "피고인 측은 공소 제기 절차의 적법성에 의문을 제기 ... (중략) ..." } ] } ...
```

## 06 편향성 완화 모델

### 파인튜닝 | GPT-3.5 Turbo



데이터 준비

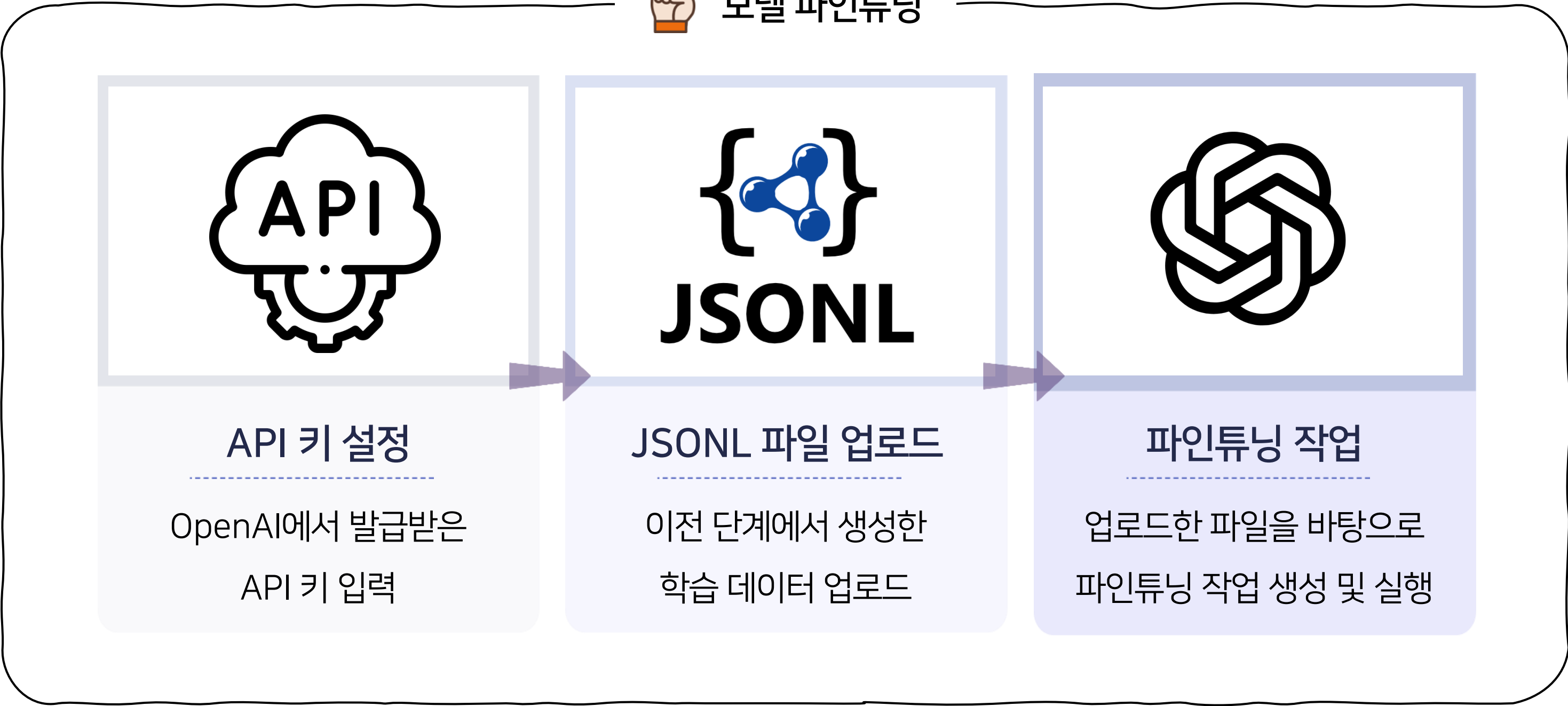
Fine-tuning용 JSONL 파일

```
{ "messages": [  
  {"role": "system", "content": "당신은 정치적으로 중립적인 뉴스로 문장을 변환하는 모델입니다." },  
  {"role": "user", "content": "다음 기사를 정치적으로 중립적인 방향으로 완화해주세요.  
[제목] : '윤 대통령 명예훼손' 언론인들 편향성이 완화된 기사를 제공  
[기사 원문] : 이 사건은 기소 자체가 위법해 공소가 기각돼야 한다고 주장 ... (중략) ...  
[편향 지수] : 0.97 (진보)}],  
  {"role": "assistant", "content": "피고인 측은 공소 제기 절차의 적법성에 의문을 제기 ... (중략) ..."} ] } ...
```

# 06 편향성 완화 모델

## 파인튜닝 | GPT-3.5 Turbo

👉 모델 파인튜닝



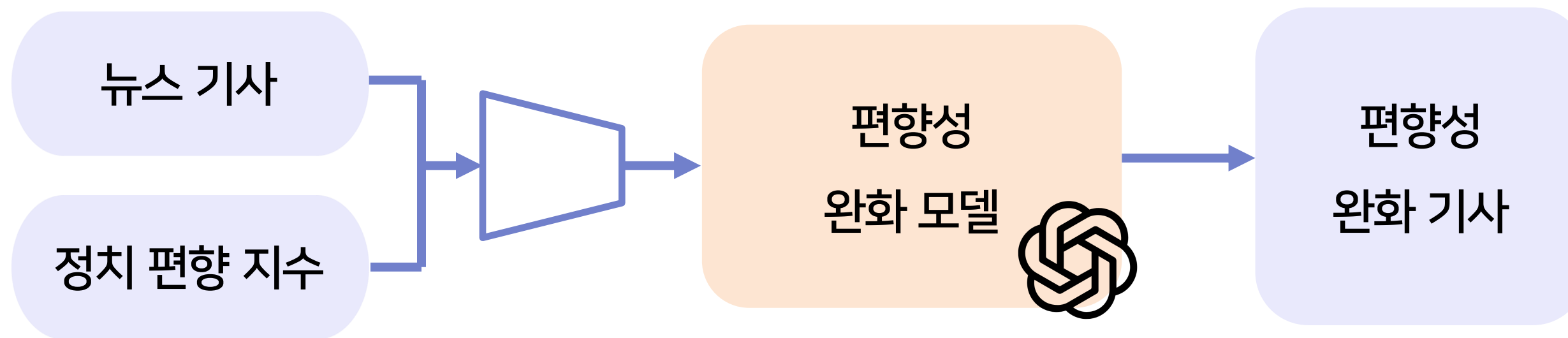
## 06 편향성 완화 모델

### 파인튜닝 | GPT-3.5 Turbo



#### 완화 작업 수행

파인튜닝 작업이 완료된 편향성 완화 모델을 사용하여  
뉴스 기사 데이터의 편향성 완화하는 작업을 수행함



앞서 산출한 **중립 지수 범위에 속한 기사**는 제외하도록 설계함

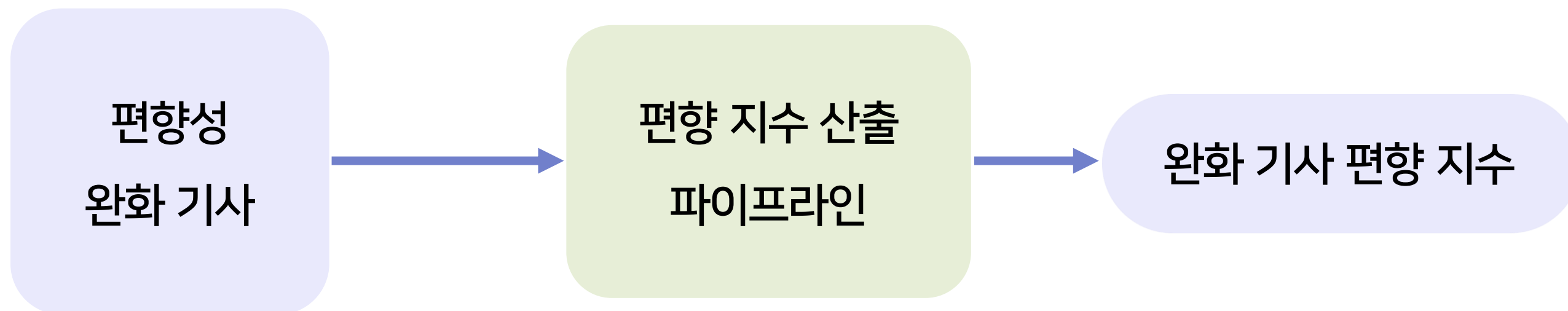


## 06 편향성 완화 모델

### 뉴스 기사 완화 결과



정치 성향 지수 산출 파이프라인을 구축하여  
완화 모델을 통해 생성된 편향성 완화 기사의 편향 지수를 측정함



# 06 편향성 완화 모델

## 뉴스 기사 완화 결과

### 최종 완화 결과

	제목	원본 기사	편향 지수	완화 기사	완화 기사 편향 지수
1	윤석열 정부 1,060일 ...3년 못 채우고 단명 자초	윤석열 전 대통령은 2년 11개월 가까 운 임기 동안 부인 김건희 여사 관련 의 혹과 여소야대 구도에서 벗어나지 못한 채 비상계엄 선포로 자멸했다 ...	0.844	윤석열 전 대통령에 대한 파면 결정으로 인해, 2022년 5월 10일 출범한 윤석 열 정부는 1,060일 만에 막을 내렸다고 합니다. 이 기간 동안 본적 기구의 활동 및 정책 추진이 이뤄졌으나 ...	0.301
...	...	...	...	...	...
75	경찰, 광화문 서십자 각 농성장서 전농 트 랙터 견인...충돌 계속	현재 비상행동 등 탄핵찬성 단체들이 도로를 점거하고 경찰과 대치하는 등 충돌하며 자하문로는 상행선과 ...	- 0.419	광화문 서십자각 농성장에서 전농 트랙 터가 경찰에 의해 견인되는 과정에서 충 돌이 발생했다. 오전 4시 15분 경 ...	- 0.009
...	...	...	...	...	...

편향 지수의 절댓값이 0.3 이상인 기사 300개를 랜덤으로 추출하여 확인한 결과,  
편향적인 기사들이 성공적으로 완화된 것이 드러남 🥰

## 06 편향성 완화 모델

### 뉴스 기사 완화 결과

#### 최종 완화 결과 예시

##### 원본 기사

윤 전 대통령은 2년 11개월 가까운 임기 동안 부인 김건희 여사 관련 의혹과 여소야대 구도에서 벗어나지 못한 채 비상계엄 선포로 **자멸했다**. 윤석열 정부는 취임과 동시에 대통령 집무실을 청와대에서 용산으로 이전했다. 건전 재정 기초를 내세우며 전임 문재인 정부와 철저히 다른 길을 가려 했다. ...

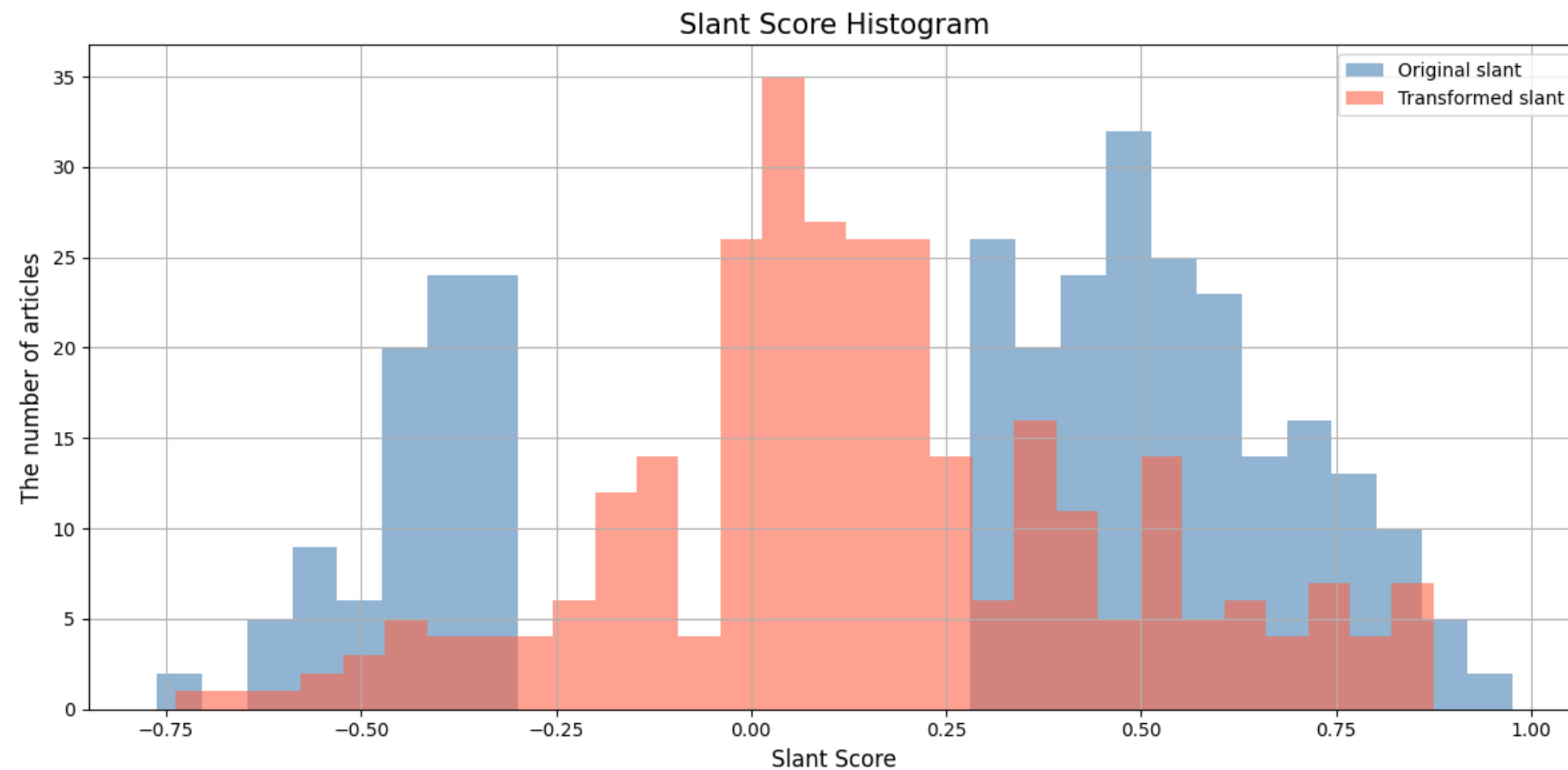
##### 원본 기사

... 윤석열 전 대통령에 대한 파면 결정으로 인해, 2022년 5월 10일 출범한 윤석열 정부는 1,060일 만에 **막을 내렸다고 합니다**. 이 기간 동안 본적 기구의 활동 및 정책 추진이 이뤄졌으나, 윤 전 대통령의 사퇴로 인해 성과 및 추후 정책 방향에 변화가 예상된다는 분석이 제기되고 있습니다. ...

실제 기사를 읽었을 때도 **정치적으로 극단적인 표현들을 완화**하여 **객관성을 확보**한 것을 확인할 수 있음

## 06 편향성 완화 모델

### 뉴스 기사 완화 결과



히스토그램으로 확인했을 때에도

편향적인 기사들이 성공적으로 완화된 것을 확인할 수 있음 🌟😊