

B반 4조 (오공이문) x 데이터청년캠퍼스

# 자연어 처리를 활용한 불공정 이용 약관 판별 서비스

## CONTENTS

- 01 개발 배경
- 02 활용 데이터 정의
- 03 데이터 처리방안 및 활용 기법
- 04 분석/개발 결과
- 05 서비스 활용방안
- 06 기대효과/향후과제
- 07 부록

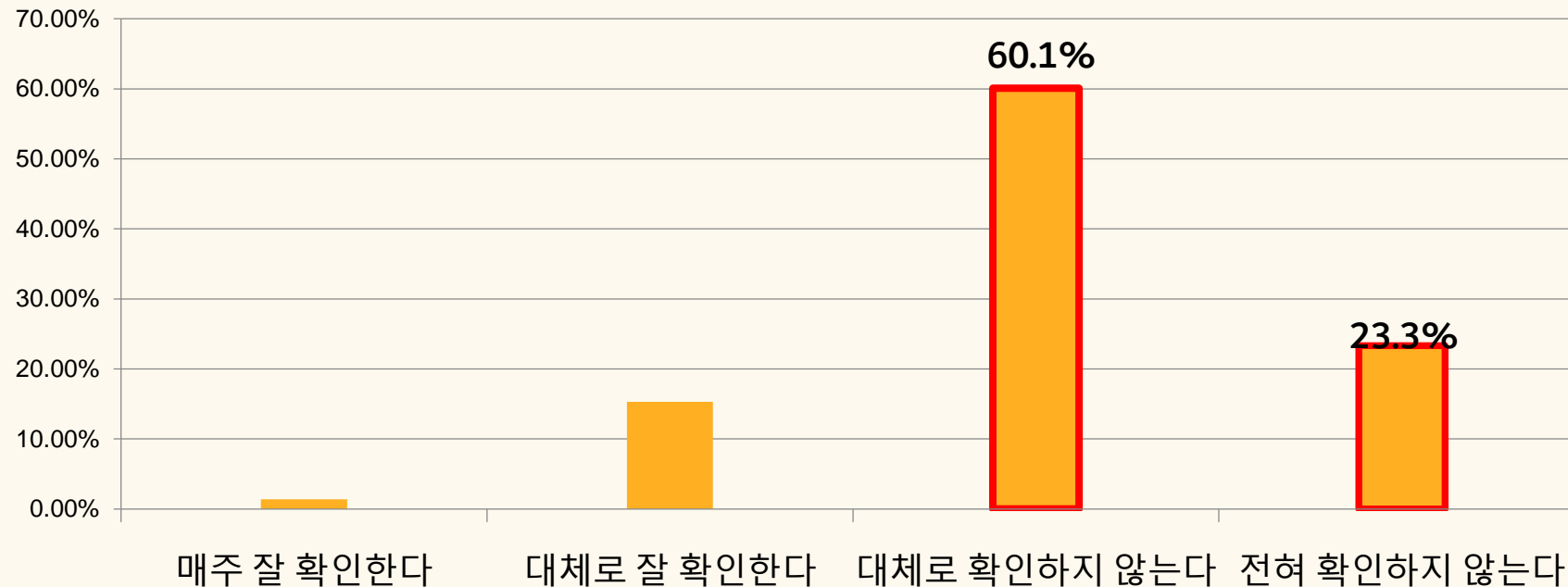
# 1 개발 배경

## 약관의 정의

약관이란 그 명칭이나 형태 또는 범위에 상관없이 계약의 한 쪽 당사자가 여러 명의 상대방과 계약을 체결하기 위하여 일정한 형식으로 미리 마련한 계약의 내용

## 이용자들의 약관 확인율

출처 : 보호위원회, 안전행정부

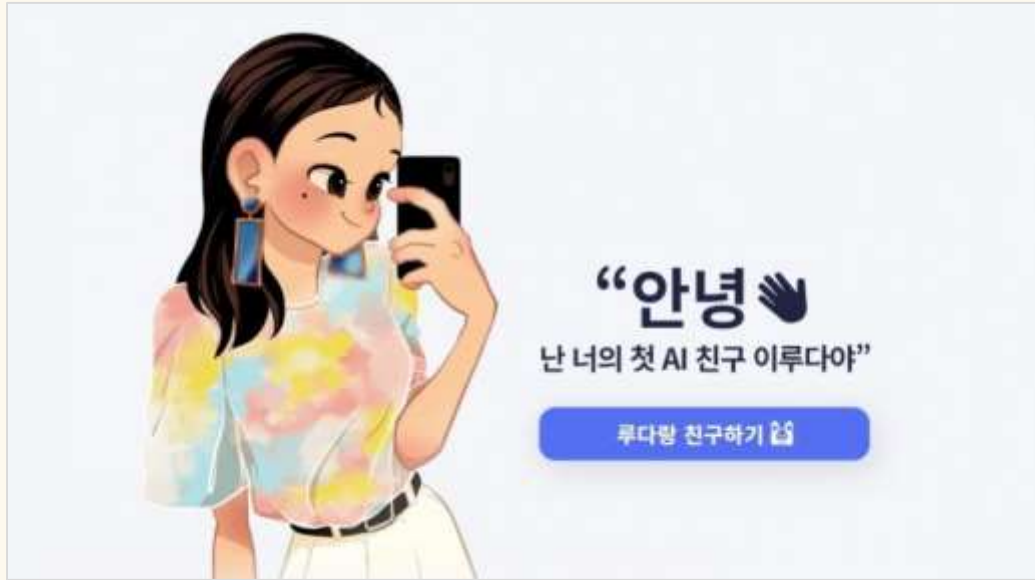


인터넷 이용자 중  
**10명 중 8명 이상**이 약관  
을 읽어보지도 않고 동의  
하는 것으로 나타남



# 1 개발 배경

## 불공정 이용약관 피해사례 1



올해 초 AI 챗봇 '이루다'는 연애정보 앱 이용자 60만여 명의 대화기록 일부와 여기에 담긴 사용자 이름, 직장 등을 무단으로 가져다 활용해 출시한 달이 안 돼 서비스를 중단했다.

## 불공정 이용약관 피해사례 2



세계 1위 전기차 업체 테슬라는 차량 결함 등 문제가 발생해도 최대 10만원만 배상하고, 고객이 차를 늦게 인수하면 차가 파손돼도 배상하지 않는다는 내용 등의 문제 조항이 공정거래위원회의 조사로 밝혀졌다.

## 데이터 셋 수집 방향

## [공정거래위원회] 4401개 데이터 셋 확보


- 불공정 약관 조항 시정 보도자료
- 불공정약관조항에 대한 의결서

## [일반이용약관] 1263개 데이터 셋 확보

- 게임
- 스트리밍
- 온라인음원제공사업자
- 웹툰
- 웹포털
- 통신사

## 2 활용 데이터 정의

### 불공정 이용 약관 시정 전/후 예시



불공정 약관 시정내용 01


**시정 전**

배달의 민족은 소비자나 음식점이 게시한 정보의 신뢰도·상품의 품질 등에 대해 어떠한 책임도 지지 않으며, 고의·중과실이 없는 한 손해배상책임을 부담하지 않는다.

배달의 민족이 소비자와 음식물을 직접 거래하지 않는 플랫폼 사업자라고 하더라도 거래과정에서 귀책사유가 있다면 그에 따른 법률상 책임이 면제될 수 없음

• **시정 후** •

음식점주 및 소비자의 귀책사유로 인해 손해가 발생해도 **배달의 민족에 고의·과실이 있다면 책임지도록 시정**하였다.



불공정 약관 시정내용 02

**시정 전**

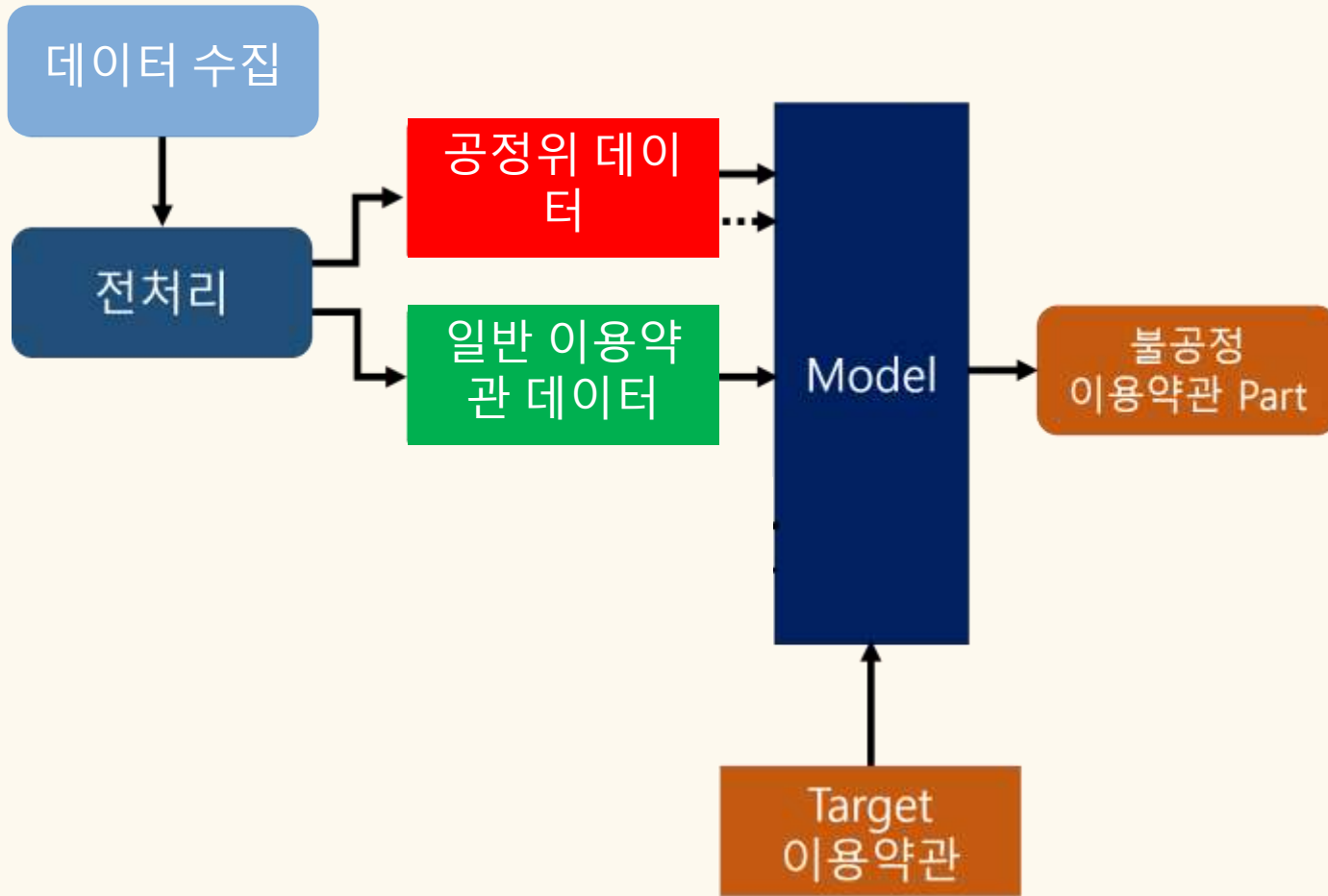
계약해지 시 소비자에게 사전에 알리는 절차를 두지 않았으며, 계약해지의 의사를 배달의 민족이 통지하기만 하면 효력이 발생하는 것으로 정하였다.

계약해지와 같이 권리의무에 중요한 영향을 미치는 행위는 사전에 통지하고 이의제기 절차를 보장하여야 하며, 민법에서 정한 의사표시 도달주의 원칙에 반해서는 안됨

• **시정 후** •

문제되는 조항을 삭제하였고, **계약을 해지하기 전 이용자에게 사전통지 절차를 보장**하는 조항을 추가하였다.

### 3 데이터 처리방안 및 활용 기법



#### 시스템 흐름 안내

1. 공정거래위원회에서 얻은 불공정 이용약관 시정 전/후 데이터와 기업 공식 홈페이지에서 찾은 일반 이용약관 데이터 셋 수집

### 3 데이터 처리방안 및 활용 기법

[활용 데이터 정의 및 예시]

의결/보도 날짜

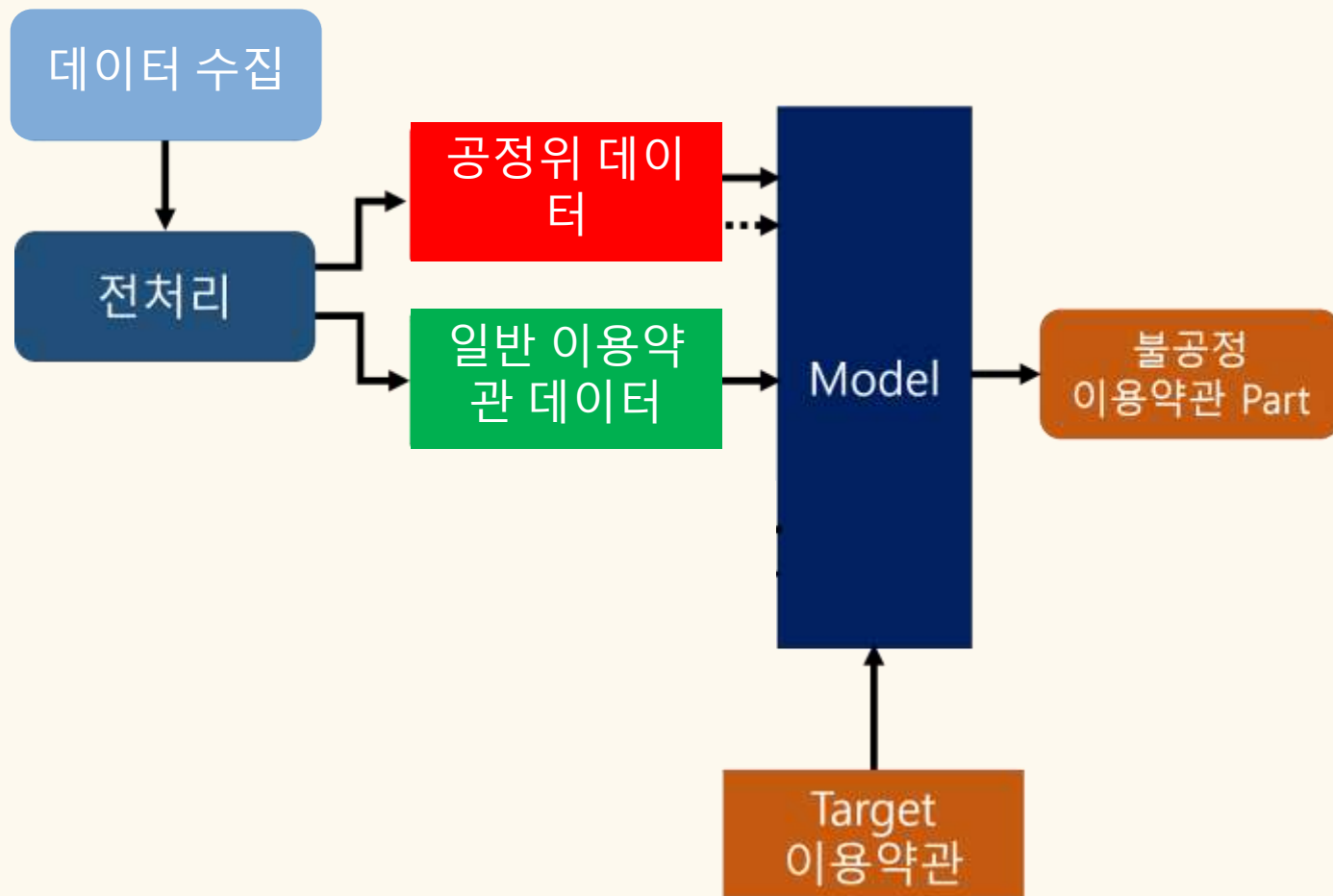
조항 내용

시정 여부  
(TRUE/FALSE)

	date	company	chapter	article	text	unfair
1	2014.03.05	LG유플러스	보상 및 환불 불가 조항	제39조 (결제취 소 및 해지)	회원은 언제든지 회원탈퇴를 하거나 뮤직 요금상품을 해지할 수 있습니다. 회원이 요금상품 해지를 신청한 경우 회사는 회원의 신청을 접수한 후 회원의 요금상품 이용기간이 끝나는 일자에 해지 처리합니다. 단, 해지신청일 기준으로 일할 계산에 의한 환불은 되지 않습니다.	TRUE
2	2014.03.05	KT 올레마켓	서비스 관련 사업자 면책 조항	제17조(서비스 기간과 중단)	회사는 본 서비스(혹은 그 일부)를 통지하거나 또는 통지함이 없이, 일시 또는 영구적으로 수정하거나 중단할 수 있으며, 본 조의 서비스의 변경, 중지로 발생하는 문제에 대해서는 어떠한 책임도 지지 않습니다.	TRUE
3	2014.03.05	KT 올레마켓	서비스 관련 사업자 면책 조항	제18조(서비스 기간과 중단)	회사는 상당한 이유가 있는 경우에 운영상, 기술상의 필요에 따라 서비스를 수정할 수 있으며, 서비스 수정을 하는 경우에는 수정 후 회사의 서비스 등을 통하여 공지합니다.	FALSE



### 3 데이터 처리방안 및 활용 기법



#### 시스템 흐름 안내

1. 공정거래위원회에서 얻은 불공정 이용약관 시정 전/후 데이터와 기업 공식 홈페이지에서 찾은 일반 이용약관 데이터 셋 수집
2. 불용어 처리, 특수기호 제거, 데이터 증강 등의 데이터 전처리 진행

### 3 데이터 처리방안 및 활용 기법

#### [핵심 구현 이론]

##### 불용어(stopword) 제거

유의미한 단어 토큰만을 선별하기 위해  
서는 불용어 제거 과정이 필요함

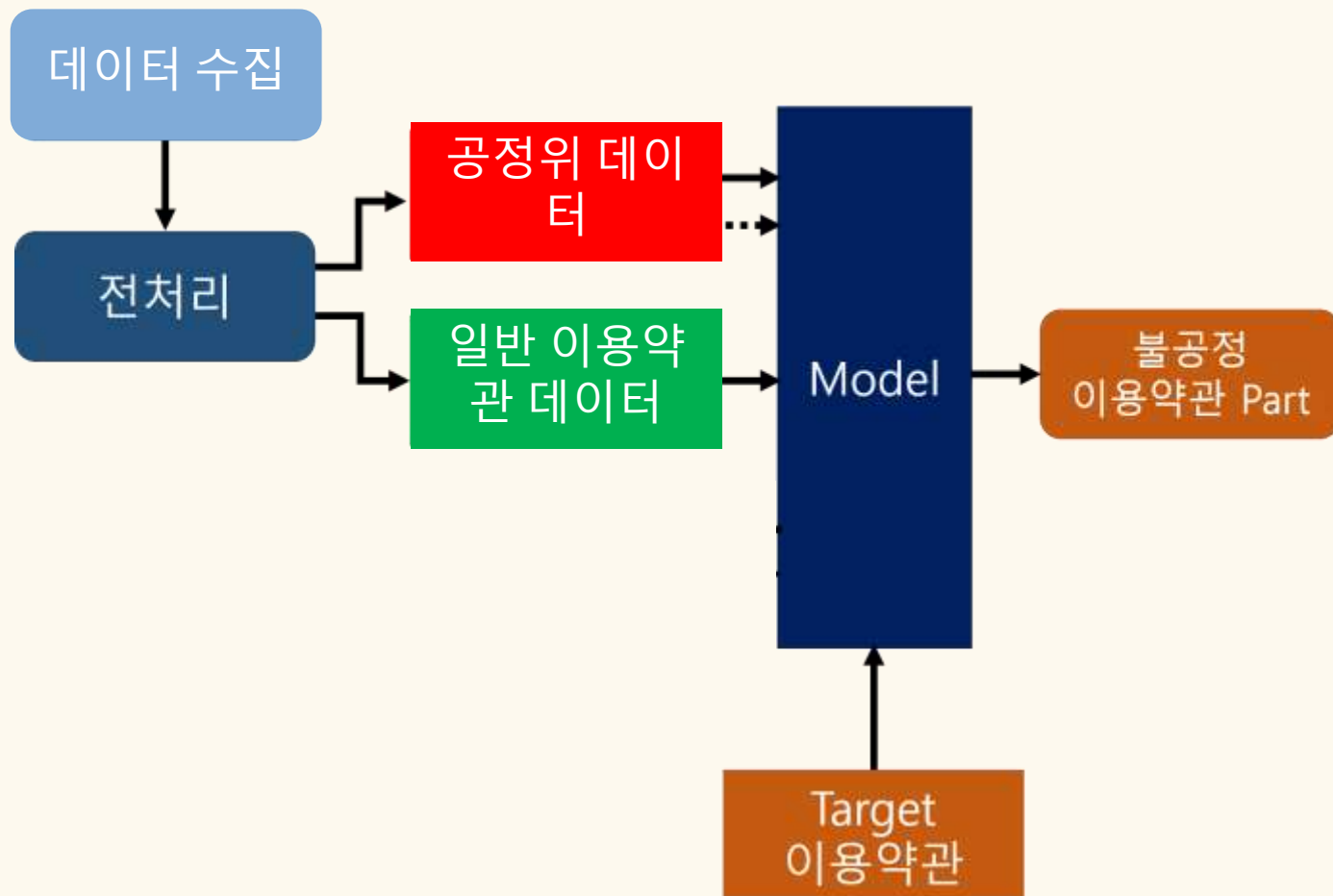
- '뭐', '으면', '을', '의' 등 27개의 불용어를 지정하고 제거

##### 데이터 증강(Data Augmentation)

학습 데이터가 부족한 상황에서 데이터를 변형  
해 그 양을 인위적으로 늘리는 기법

- RS(Random Swap)  
: 문장 내 임의의 두 단어의 위치를 바꿈
- RD(Random Deletion)  
: 임의의 단어를 삭제

### 3 데이터 처리방안 및 활용 기법



#### 시스템 흐름 안내

1. 공정거래위원회에서 얻은 불공정 이용약관 시정 전/후 데이터와 기업 공식 홈페이지에서 찾은 일반 이용약관 데이터 셋 수집
2. 불용어 처리, 특수기호 제거, 데이터 증강 등의 데이터 전처리 진행
3. 한국어 특화 BERT모델인 KoBERT모델을 supervised learning 진행

### KoBERT

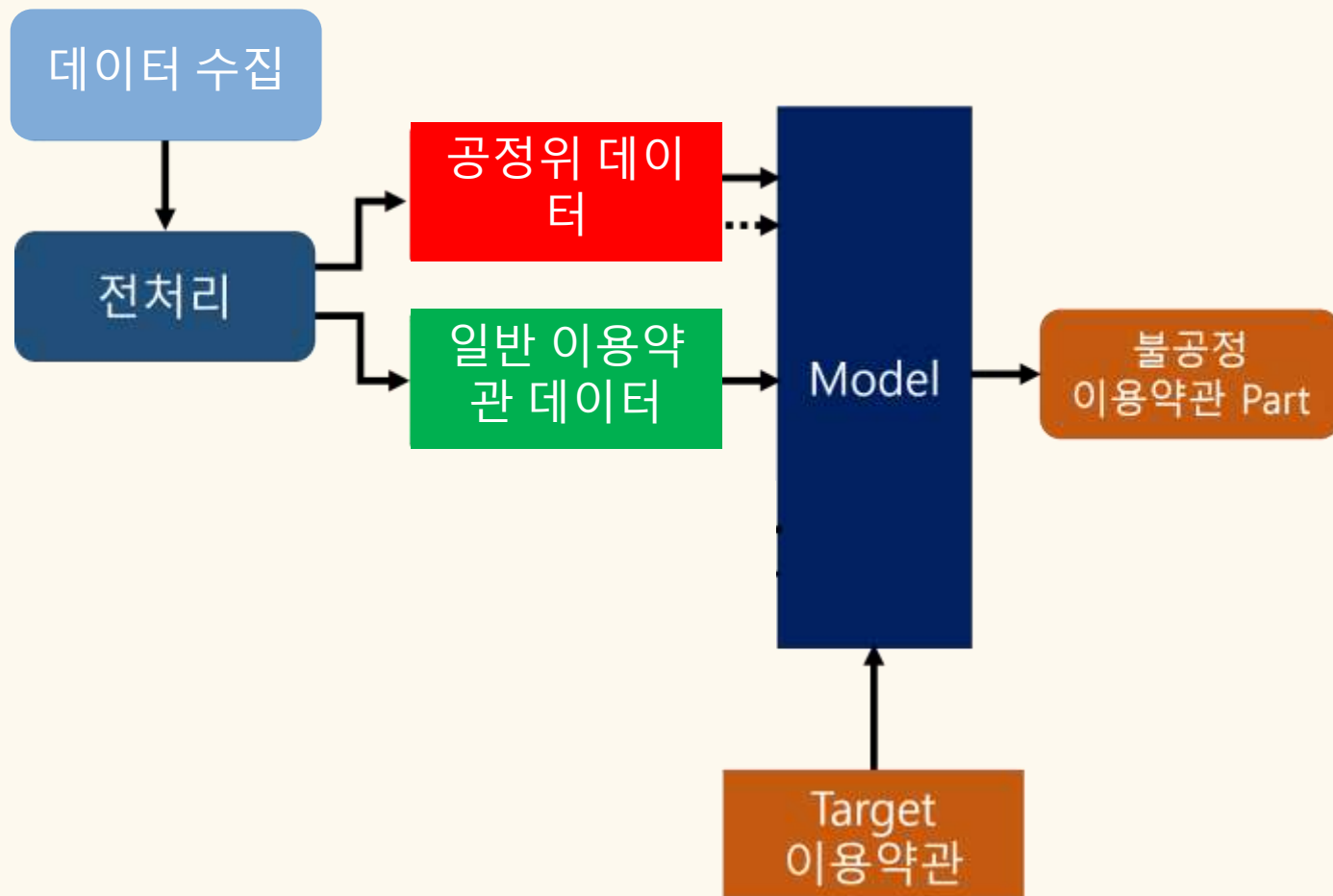
- BERT 모델에서 한국어 데이터를 추가로 학습시킨 모델로, BERT 모델을 한국어에도 잘 활용할 수 있도록 만들어진 것
- 한국어 위키에서 5백 만 개의 문장과 54백 만 개의 단어를 학습한 모델

### 지도 학습(Supervised Learning)

입력 값이 주어지면 입력 값에 대한 Label를 주어 학습시키는 것.

- 분류(classification)  
: 주어진 데이터를 정해진 라벨에 따라 분류하는 문제

### 3 데이터 처리방안 및 활용 기법



#### 시스템 흐름 안내

1. 공정거래위원회에서 얻은 불공정 이용약관 시정 전/후 데이터와 기업 공식 홈페이지에서 찾은 일반 이용약관 데이터 셋 수집
2. 불용어 처리, 특수기호 제거, 데이터 증강 등의 데이터 전처리 진행
3. 한국어 특화 BERT모델인 KoBERT모델을 supervised learning 진행
4. 학습된 모델에 target 이용약관을 넣어 불공정 이용약관 part 하이라이트 및 위험도 표시

## 구글 크롬 확장 프로그램(google chrome extension)

- 구글 크롬 브라우저를 수정하는 브라우저 확장 프로그램
- HTML, JavaScript 및 CSS와 같은 웹 기술을 기반으로 작성
- 예시) Grammarly

: 무료로 영어 문법을 교정 및 검사

**Tooltip Dictionary**

: 찾고자 하는 단어에 마우스를 올리면 툴팁으로 사전 제공



## 모델 학습 결과

```

/usr/local/lib/python3.7/dist-packages/torch/utils/data/dataloader.py:481: UserWarning: This DataLoader
  cpuset_checked))
epoch 1 batch id 1 loss 0.7111360430717468 train acc 0.515625
epoch 1 train acc 0.7028138102725366
100% ██████████ 27/27 [00:13<00:00, 2.43it/s]
epoch 1 test acc 0.8663194444444444
100% ██████████ 106/106 [02:25<00:00, 1.22s/it]
epoch 2 batch id 1 loss 0.43623605370521545 train acc 0.78125
epoch 2 train acc 0.8917551100628931
100% ██████████ 27/27 [00:13<00:00, 2.39it/s]
epoch 2 test acc 0.9261111111111111
100% ██████████ 106/106 [02:26<00:00, 1.23s/it]
epoch 3 batch id 1 loss 0.18622085452079773 train acc 0.90625
epoch 3 train acc 0.9380896226415094
100% ██████████ 27/27 [00:13<00:00, 2.42it/s]
epoch 3 test acc 0.9263657407407407
100% ██████████ 106/106 [02:26<00:00, 1.23s/it]
epoch 4 batch id 1 loss 0.07569917291402817 train acc 0.953125
epoch 4 train acc 0.9681931341719078
100% ██████████ 27/27 [00:13<00:00, 2.40it/s]
epoch 4 test acc 0.9313194444444444
100% ██████████ 106/106 [02:26<00:00, 1.23s/it]
epoch 5 batch id 1 loss 0.024801865220069885 train acc 1.0
epoch 5 train acc 0.9839655398322852
100% ██████████ 27/27 [00:13<00:00, 2.43it/s]
epoch 5 test acc 0.9318981481481481

```

모델 학습 결과 93.18%의  
정확도를 보였다.

## 4 분석/개발 결과

### 모델 학습 결과

(불공정 TRUE 데이터)

약관 입력 : 甲은 회원가입일로부터 발생시킨 당월 통화료를 익월 25일까지 납부하지 아니할 경우 일시 정지와 함께 직권해지되(될 수 있으며)며 이 경우 乙은 甲에게 고지의 의무를 갖지 아니한다.  
[0.04841921 0.9416199]

94.2%의 확률로 불공정 약관일 가능성이 있습니다.

**한국인터넷텔레콤의 회원가입약관상 불공정약관조항 판별 결과: 94.2%의 확률로 불공정 약관 판단**

약관 입력 : 본 계약에 명시되지 않은 사항과 해석상 차이가 있을 때는 갑,을 쌍방 합의에 의하여 처리하되 합의가 이루어지지 않을 때에는 사업협회의 결정에 따른다.  
[0.24801935 0.75322866]

75.3%의 확률로 불공정 약관일 가능성이 있습니다.

**한국트라의 위.수탁 관리계약서 및 합의각서상 불공정약관조항 판별 결과: 88.1%의 확률로 불공정 약관 판단**



## 4 분석/개발 결과

### 모델 학습 결과

(불공정 FALSE 데이터)

약관 입력 : “요기요” 는 이용자가 구매한 재화 등이 시간을 초과하여 배달되거나 약속한 시간 내에 배달되지 않음으로 인한 이용자의 손해에 대하여 요기요의 귀책사유에 따른 책임을 부담할 수 있습니다.  
[0.8174905 0.1553236]  
불공정 약관이 아닙니다.

배달 관련 조항 판별 결과: 불공정 약관 x

약관 입력 : 회사” 는 다음과 같은 사유가 있는 경우, 이용계약을 해지할 수 있습니다. 다. 기타 “회원” 이 이 약관에 위배되는 행위를 하거나 이 약관에서 정한 해지사유가 발생한 경우  
[0.6405901 0.4149038]  
불공정 약관이 아닙니다.

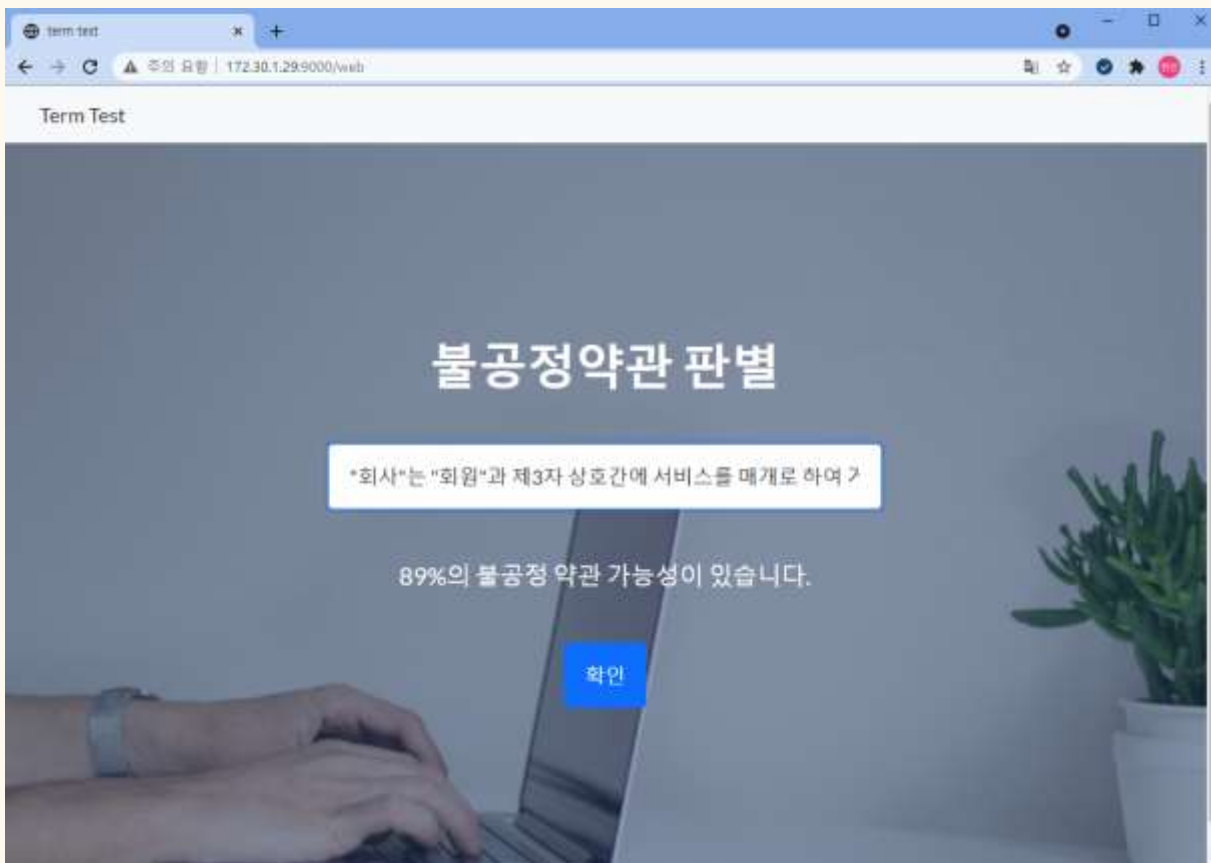
이용 계약의 종료 조항 판별 결과: 불공정 약관 x

약관 입력 : 회사는 귀하가 작성하는 리뷰 등 게시물을 소중하게 생각하며 변조, 훼손, 삭제되지 않도록 최선을 다하여 보호합니다.  
[0.92368025 0.07232389]  
불공정 약관이 아닙니다.

게시물 관련 조항 판별 결과: 불공정 약관 x

## 5 서비스 활용방안

### 웹 사이트 이미지



### Chrome extension 활용 서비스 구현 이미지

랭킹
만남
썸
연애

2. 유변상태의 계성으로는 로그인 이용이 제한되고, "회사"는 유변상태를 제공하지 않습니다.

3. "회사"는 제1항의 처리예정일로부터 30일 이전에 해당 사실의 발생을 "회원"에게 통지하고, "회원"은 별도의 인증 절차를 통해 휴면을 해제할 수 있습니다.

### 제 22 조 (책임제한)

1. "회사"는 천재지변, 디도스(DDOS)공격, IDC장애, 기간통신사업자의 정당한 서비스 제공을 중단하는 경우 "회사"는 "서비스"를 제공할 수 없는 경우에는 "서비스" 제공을 중단할 수 있습니다.

2. "회사"는 "회원"의 귀책사유로 인한 "서비스" 이용의 장애에 대하여는 책임을 지지 않습니다.

3. "회사"는 "서비스"에 게재된 "콘텐츠"의 정보, 자료, 사실의 신뢰도, 정확성, 적법성 등에 관하여는 책임을 지지 않습니다.

4. "회사"는 무료로 제공되는 "서비스" 이용과 관련하여 관련 법령에 특별한 규정이 없는 한 책임을 지지 않습니다.

5. "회사"는 "회원"과 제3자 상호간에 서비스를 매개로 하여 거래 등을 한 경우에는 책임이 면제됩니다.

6. "회사"는 "회원"의 모바일 환경이나 "회사"의 관리 범위에 있지 아니한 보안 문제로 인하여 발생하는 제반 문제 또는 현재의 보안기술 수준으로 방어가 곤란한 네트워크 해킹 등 "회사"의 귀책사유 없이 발생하는 문제에 대해서 책임을 지지 않습니다.

7. "회원"이 자신의 개인정보를 타인에게 유출 또는 제공함으로써 발생하는 피해에 대해서는 "회사"는 일체의 책임을 지지 않습니다.

8. "회사"는 "회원"이 "서비스"를 이용하여 기대하는 효용을 얻지 못한 것에 대하여 책임을 지지 않으며 "서비스"에 대한 취소 선택 또는 이용으로 발생하는 손해 등에 대해서는 책임이 면제됩니다.

9. WiFi환경이 아닌 곳에서 접속하여 데이터 요금 발생 시, "회원"은 부과된 요금에 대해 "회사"에게 일체 지불 요청을 할 수 없습니다.

10. "회사"는 "서비스" 개편이나 운영상 또는 "회사"의 긴박한 상황 등에 의해 "서비스" 전부나 일부를 중단할 필요가 있는 경우 30일 전에 이를 공지하고 "서비스"의 제공을 중단할 수 있습니다. "회원"은 "서비스" 전부 혹은 일부의 종료 시 사용 기간이 남아 있지 않은 정액제 아이템 또는 횡수제 아이템에 대하여 배상을 청구할 수 없습니다.

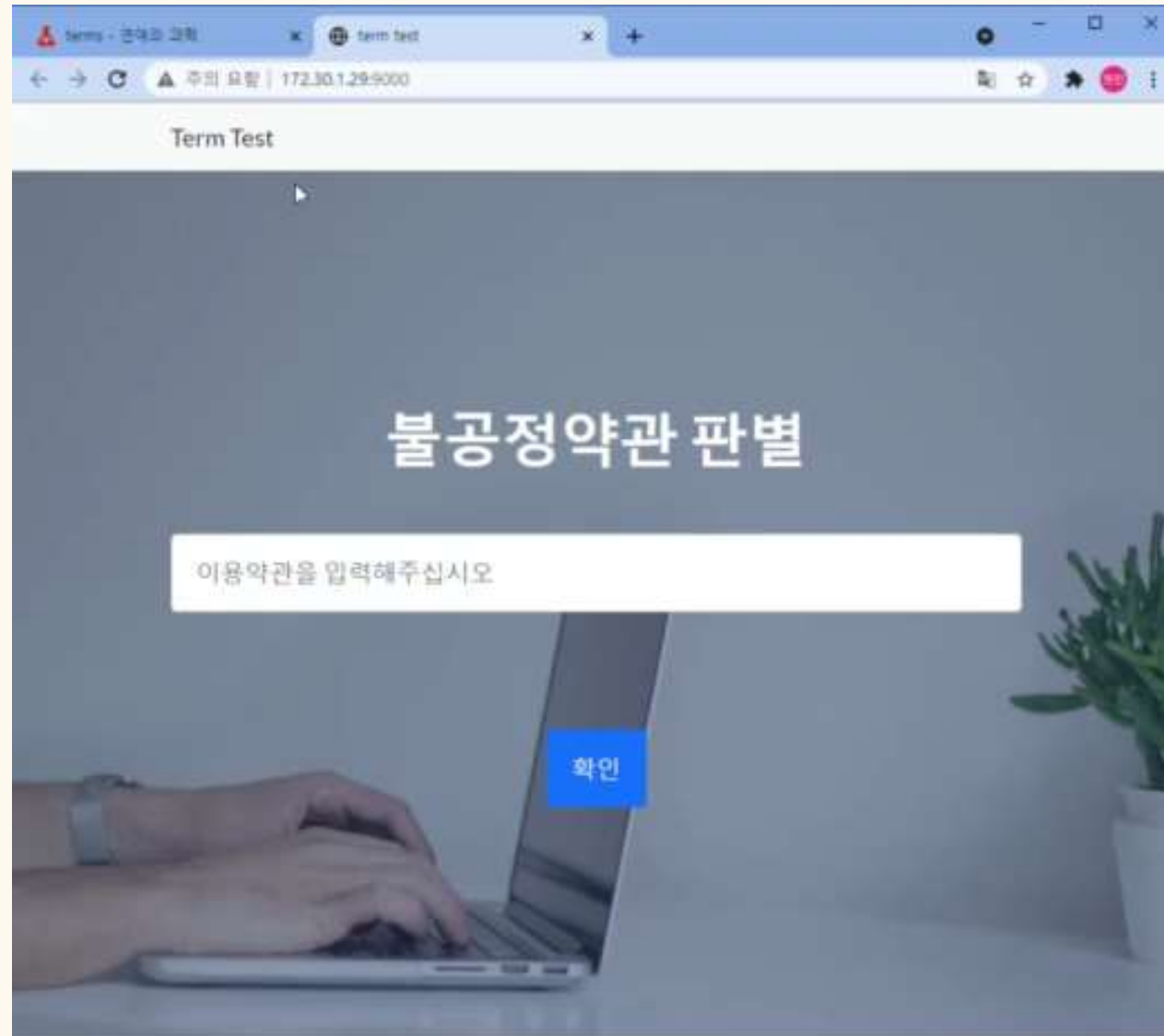
**불공정약관 판별**

"회사"는 "회원"과 제3자 상호간에 서비스를 매개로 하여 거래 등을 한 경우에는 책임이 면제됩니다.

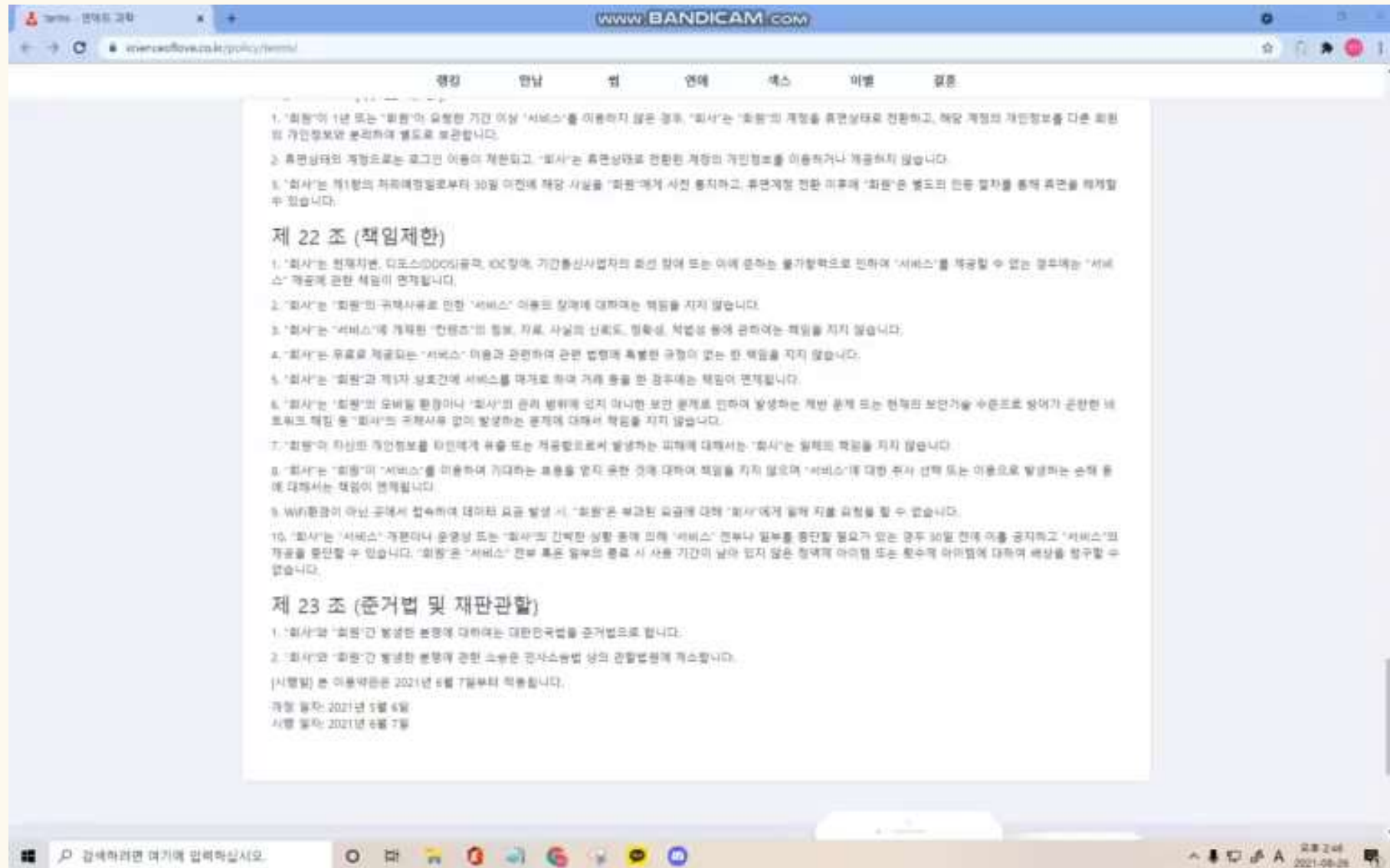
확인

89%의 불공정 약관 가능성이 있습니다.

웹 사이트를 통한 불공정 이용약관 판별 영상



## Chrome Extension을 통한 불공정 이용약관 판별 영상



### 차별점 & 한계점



## 6 기대효과

약관은 사업자가 일방적으로 마련한 계약

→ 소비자의 정상적인 기대나 합리적 이익을 약관에 제대로 반영하지 않을 수 있음

약관 동의 시 한 번 더  
검토해 볼 수 있다

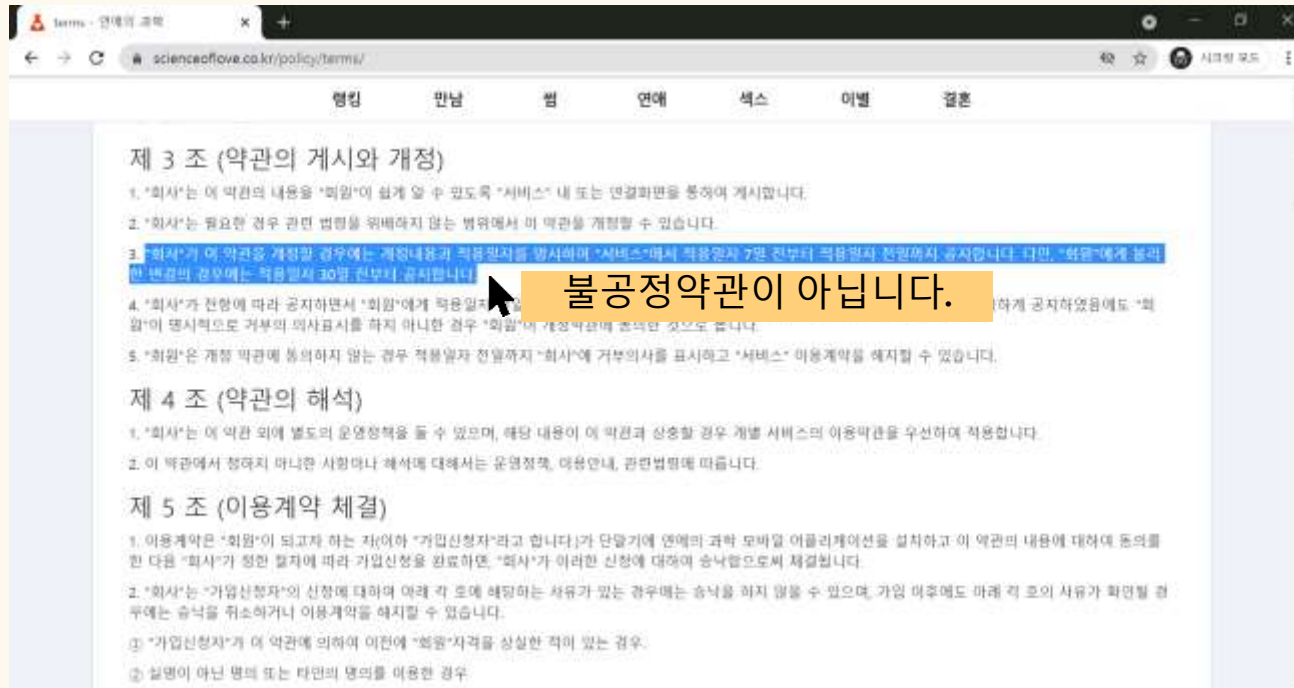
약관 동의에 대한  
책임감을 부여한다

불공정 약관으로부터의  
피해를 예방한다

➔ 기대효과를 통하여 부가적인 사회적/경제적 가치가 창출 될 것으로 기대된다.

## 6 향후 과제

### ① 웹사이트 개선



마우스 커서 블록 처리를 통해 불공정 약관 여부를 판별한다.  
또한, 외부접근을 위해 포트 포워딩을 통해 웹을 배포하도록 한다.

### ② 개인정보이용약관 등 약관의 범위 확대

개인정보 이용약관 중심

계약 검토 중심

금융 거래 약관 중심

기존의 서비스 이용약관 뿐만 아니라 범위를 넓혀 개인정보 이용약관, 계약 검토, 금융거래 약관 등으로 넓혀나갈 것이다.

## 7 부록(참고문헌)

### [논문]

- 조용현.차영균, 딥러닝을 활용한 개인정보 처리방침 분석 기법 연구, 정보보호학회논문지, 305-312, 2020
- Jason Wei·Kai Zou, EDA : Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks, EMNLP-IJCNLP 2019 short paper, 2019
- 황상흠.김도현, 한국어 기술문서 분석을 위한 BERT 기반의 분류모델, 한국전자거래학회지, 203-214, 2020
- 윤현식.윤영호.박현재, 머신러닝을 통한 SNS상의 성매매 알선 홍보 글 탐지의 효율성 제고방안 연구, 한국지 역정보화학회지, 43-65, 2020
- KoBERT <https://github.com/SKTBrain/KoBERT>
- [사외투]사전 및 mecab 적용 : <https://somjang.tistory.com/entry/Google-Colab%EC%97%90%EC%84%9C-mecab-ko-dic-%EC%82%AC%EC%9A%A9%EC%9E%90-%EC%82%AC%EC%A0%84-%EC%B6%94%EA%B0%80%ED%95%98%EA%B8%B0>
- naver\_review\_classifications\_pytorch\_kobert : [https://colab.research.google.com/github/SKTBrain/KoBERT/blob/master/scripts/NSMC/naver\\_review\\_classifications\\_pytorch\\_kobert.ipynb](https://colab.research.google.com/github/SKTBrain/KoBERT/blob/master/scripts/NSMC/naver_review_classifications_pytorch_kobert.ipynb)
- KorEDA : <https://github.com/catSirup/KorEDA/tree/master>



## [불공정 이용약관

## 관] 공정거래위원회 이용약관관 의결서

<https://case.ftc.go.kr/ocp/co/ltfr.do?pageIndex=1&caseNo=&caseNm=%EC%9D%B4%EC%9A%A9%EC%95%BD%EA%B4%80&decsnNo=&startRceptDt=&endRceptDt=&reprsntManagtTyCd=&reprsntVioltTy=&searchKrwrd=#link>

- 공정거래위원회 보도자료 <https://ftc.go.kr/www/ReportUserList.do?key=125&rpttype=>

## [일반 이용약관관] (각 회사 홈페이지 참고)

- 게임 : 넥슨, 피망, 넷마블, 카카오게임즈, **NHN**, 스마일게이트알피지, 데브시스터즈, 액토즈소프트
- 스트리밍 : 웨이브, 왓챠, 티빙, 넷플릭스, 시즌, **LGU+모바일TV**, 아프리카**TV**, **twitch**(트위치), **VLIVE**
- 온라인음원제공사업자 : 멜론, 지니뮤직, **flow**
- 웹툰 : 네이버웹툰, 레진코믹스?, 리디북스, 탑툰
- 웹포털 : 네이버, 카카오
- 통신사 : **KT**, **SKT**, **U+**

B반 4조 (오공이문) x 데이터청년캠퍼스

감사합니다.