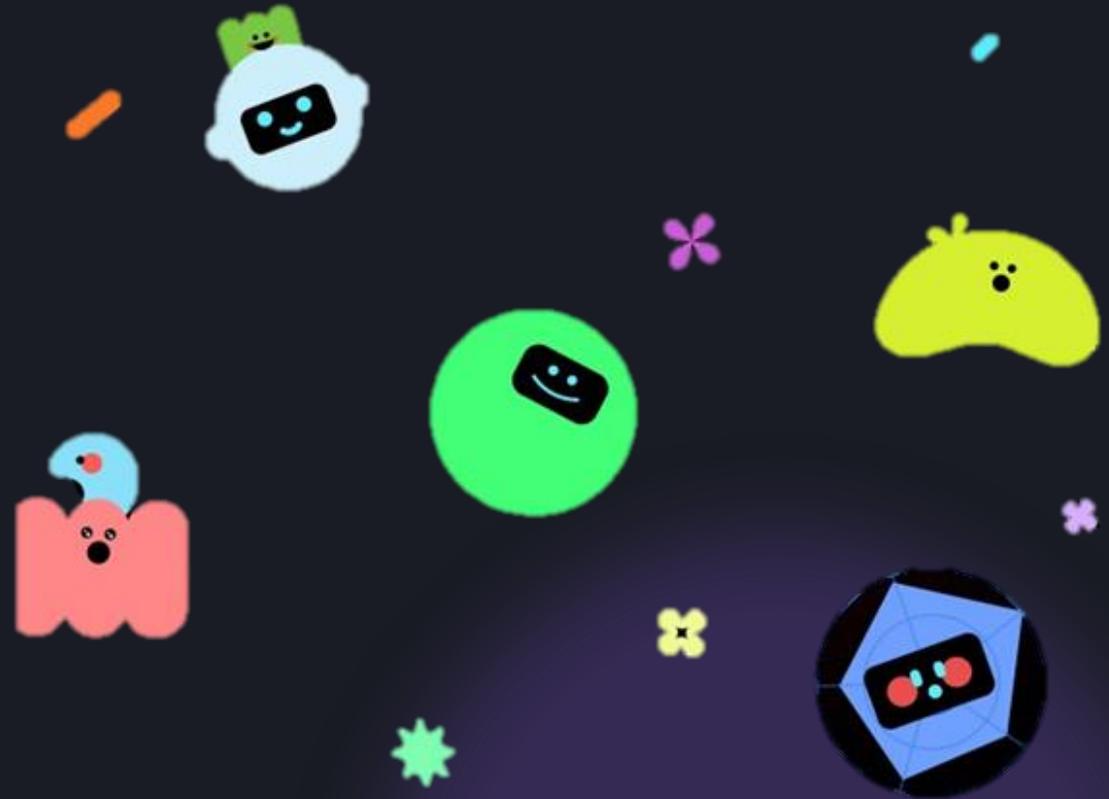


# 낙시 기사 여부 & 유형 판별 서비스 PRESSPULSE

5조 낭만어부

BANDALCOM, SR.Lee, HJ.Kang,  
SH.Jeon, ML.Heo, CM.Lee



# 목차

## 01 개요

R&R  
주제 선정 배경  
PressPulse 차별점 목표

## 02 구축

파이프라인 설계  
데이터수집  
EDA 및 전처리  
학습 모델 훈련  
기능확장  
배포 / 데모

## 03 결론 및 활용 방안

결론  
PressPulse 확장성 제안

# 01. 개요

# R&R



BANDALCOM

Data Preprocess  
Data-Centric  
모델 훈련

HJ.Kang

LLM 연동  
세부 프로그래밍

SR.Lee

Data Preprocess  
Model-Centric  
모델 훈련

CM.Lee

LLM 연동  
LLM 자료조사

SH.Jeon

LLM 배포  
PPT 제작

ML.Heo

LLM 배포  
서비스 구현

# 주제 선정 배경



Q. 제목만 봤을 때, 둘 중 어느 기사가 낚시성 기사일까요?  
**조회수**를 높이기 위해

1

**자극적이거나 오해의 소지가 있는** "김재준 변호사, 양자택일... 중요할까"

기사입력 2023.02.02 09:27 / 기사수정 2023.02.02 15:29

**기사 제목을 사용하는 경향**

2

'장동건♥' 고소영 "'여보 나도 할 말 있어' 이윤미 또 다른 모습"

기사입력 2023.11.02 17:03

출처: 엑스포츠 뉴스

# 주제 선정 배경



● 매우 문제가 됨 ● 약간 문제됨 ● 보통 ● 별로 문제 아님 ● 전혀 문제 안됨  
(n=1,956, 단위: %, 점, 5점 평균)

허위·조작정보(가짜뉴스) (4.36점)



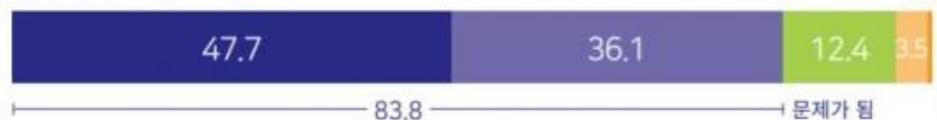
뉴스성 기사 (4.33점)



어뷰징 기사 (4.31점)



언론사의 오보 (4.27점)



'한국의 언론인 2019' 에서 '언론 보도 중 가장 심각한 문제' 항목 조사 결과 (자료=한국언론진흥재단)

사용자를 저품질 콘텐츠로 유도

이용자에게 불편함을 느끼게 할 뿐만 아니라  
제목만으로 일부 사람들에게  
그릇된 인식을 심기도 한다.

〈클릭베이트, 뉴스성 기사의 위협〉 김민 기자, 2021, 미디어 경청

# 주제 선정 배경



## Click-bait가 가져오는 사회 문제로 인해 낱시성 기사의 위험성 언급과 다양한 해결방안 개발

조선일보 · A4면 1단 · 2023.03.04. · 네이버뉴스

좋은 뉴스 사이트는 기사 오류 바로잡고 낱시성 제목 없어야

개

0기 YTN사이언스 · 2018.07.14.

현: 낱시성 기사...인공지능이 잡는다

기

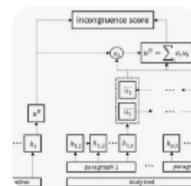
구 전자신문 · 1면 3단 · 2019.01.21. · 네이버뉴스

조

서울대·KAIST, '낱시성 기사' 잡아내는 딥러닝 모델 기술 개발

내용과 동떨어진 과장된 제목으로 네티즌의 클릭을 유도하는 '낱시성 기사'를 걸러낼 수 있는 기술이 개발됐다. 포털이나 검색 사이트 보조 시스템으로 활용할 수 있는 기술이다. KAIST는 차미영 전산학부 교수가 정교...

신뢰성	
1. 리워 정보 반복 게시 금지	배점 22
2. 책임성 있는 정보인가	18
3. 이모티콘 소정의인 고지워 거	19



www.joongang.co.kr > 사회 > 전국

[단독] 제목에 마우스 올려놓자 "낱시성 기사일 확률 88%" | 중...

2018.05.07. 또 "몸무게 54kg에 55사이즈"유지 중인 올해 '52세' 슈퍼 동

안

본문

시' MV

30일 오후 6시 각종 음원 사이트에는 박효신의 신곡 '별 시(別時)'가 공개

권길여 기자 · 2018-04-30 22:19:51

낱시성 기사일 확률 : 88%

얼굴 반쪽 돼 '리즈' 갱신한 엄지가 말한 다이어트 비법

여자친구의 엄지가 쇼케이스서 몰라보게 예뻐진 미모의 비결을 털어냈다.

황비 기자 · 2018-04-30 21:43:17



# 뉴스 기사 여부 & 유형 판별 서비스

언론의 맥박, 언론이 건강한지 체크한다

**Press**  **Pulse**

# PRESSPULSE 차별점 목표 - 서비스

## 사용자

### 선택적 정보 취득

총 10가지 유형의  
낙시성 기사 분류

### 효율적 정보 취득

기사 요약 및  
빈출 단어 제공

### 유저 친화적인 UI

사용자들의  
접근 용이

### '미디어리터러시' 능력 증진

기사를 보는  
비판적인 시각 증진

## 언론사/기자

### 저널리즘의 '진실성'

'진실성' 의무에 대한 경각심 전달

### 발행 전 재검토

오해를 불러일으키는 단어/문장 확인 가능

# PRESSPULSE 차별점 목표 - 연구

대규모 언어모델(LLM)로  
의도 분류

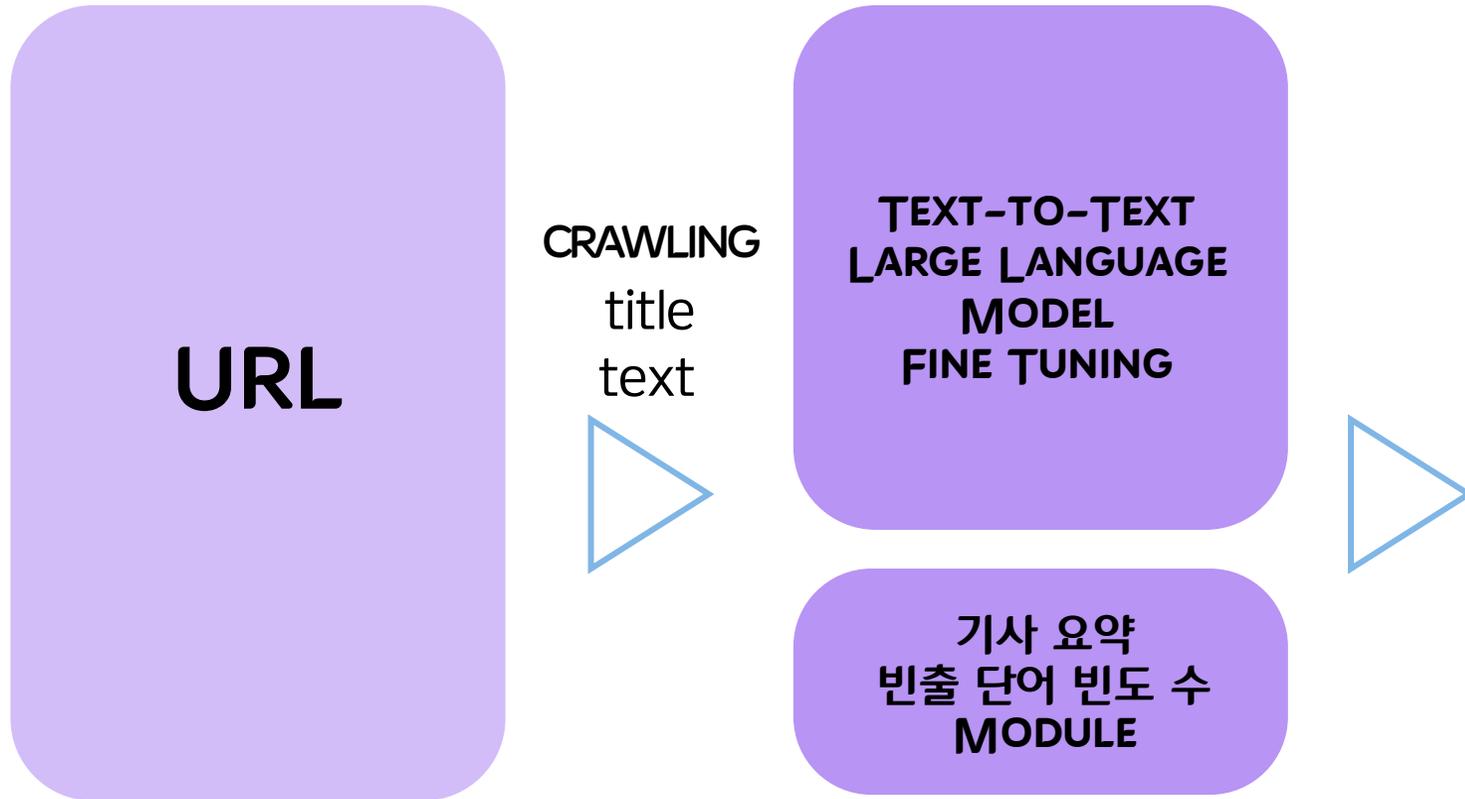
낱시성 기사 판별  
&  
유형 분류 서비스  
제공

Model & Data-Centric  
Perspective

모델의  
Hyperparameter 변경  
&  
데이터를 다양하게  
변경하며 접근

## 02. 구축

# 파이프라인 설계



진실을 찾아가는 PressPulse, 언론의 건강상태를 체크하세요.

기사 URL을 입력하면 낚시 여부를 판별해주는 시스템입니다.

1. 기사 URL을 입력하시고 'submit' 버튼을 눌러주세요
2. 낚시기사 판별 여부, 낚시 유형을 확인 가능합니다 (낚시기사가 아닐 시 유형이 나오지 않습니다)

url

URL을 입력해주세요

Clear Submit

낚시성 판별 결과

요약 정보

빈출 빈도 수

Flag

## 비플라이소프트, AI기반 `낚시성 기사 탐지 데이터` 구축 완료

입력: 2023-01-17 17:21 노회근 기자



TT 폰트

N 기자 구독

낚시성 기사 탐지 위한 AI 언어 모델 및 학습용데이터 72만 건 구축  
아이서퍼, 로제우스 등 미디어 특화 서비스 플랫폼 고도화

## AI 허브, 낚시성 기사 탐지 데이터

데이터 크기 : 733,427 ROWS, 15 COLUMNS

## 학습에 필요한 Feature 사용

기사 ID

newsID

기사 카테고리

newsCategory

기사 하위 카테고리

newsSubcategory

기사 제목

newsTitle

기사 부제목

newsSubTitle

기사 본문

newsContent

세부 번호

partNum

용도 유형

useType

처리 유형

processType

처리 패턴

processPattern

처리 수준

processLevel

문장 수

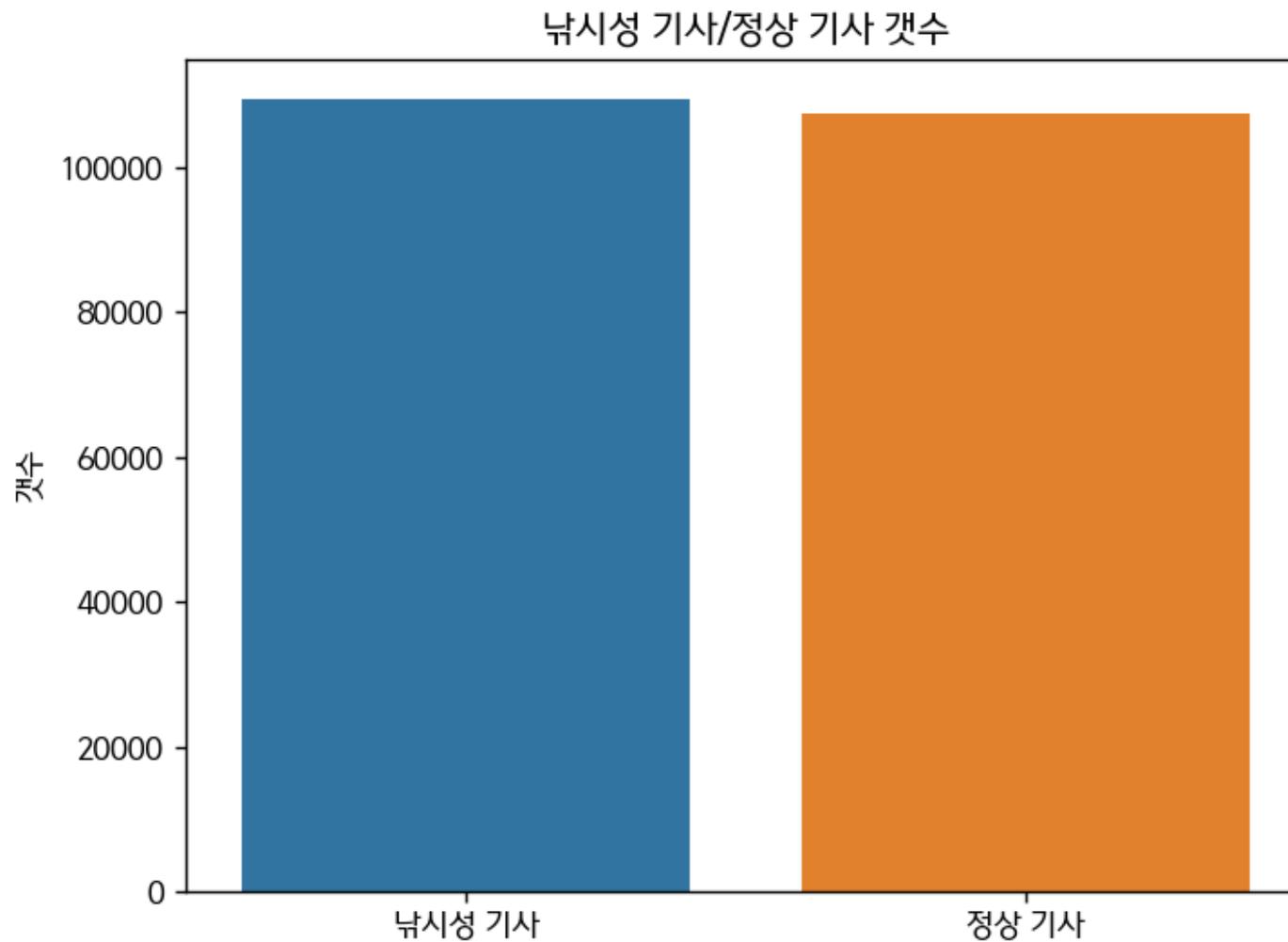
sentenceCount

기사부제목('newsSubTitle') NULL인 ROWS 삭제

낙시성기사유형('processPattern') 코드 이름 변경

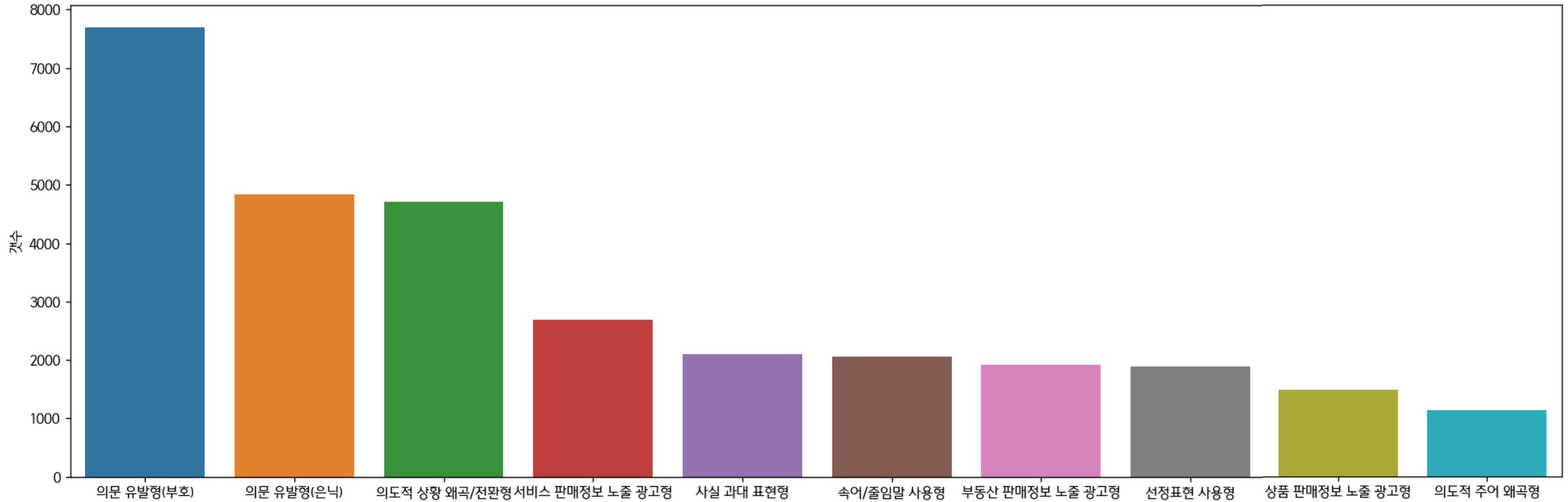
11	의문 유발형(부호)	12	의문 유발형(은닉)
13	선정표현 사용형	14	속어/줄임말 사용형
15	사실 과대 표현형	16	의도적 주어 왜곡형
21	상품 판매정보 노출 광고형	22	부동산 판매정보 노출 광고형
23	서비스 판매정보 노출 광고형	24	의도적 상황 왜곡/전환형

# EDA 및 전처리



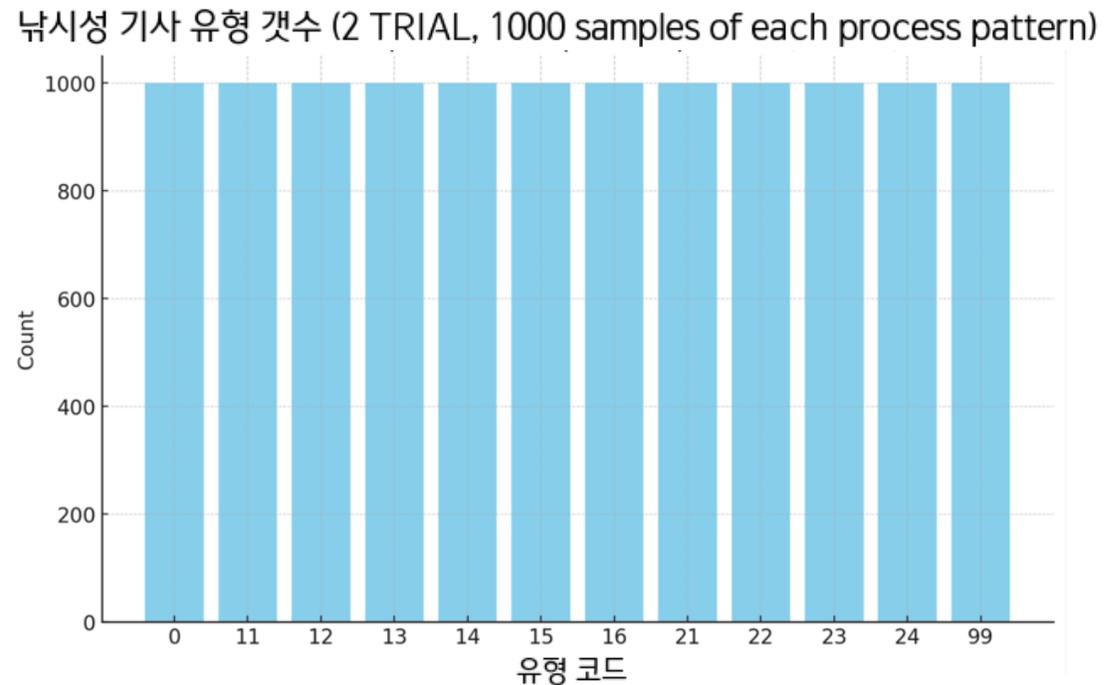
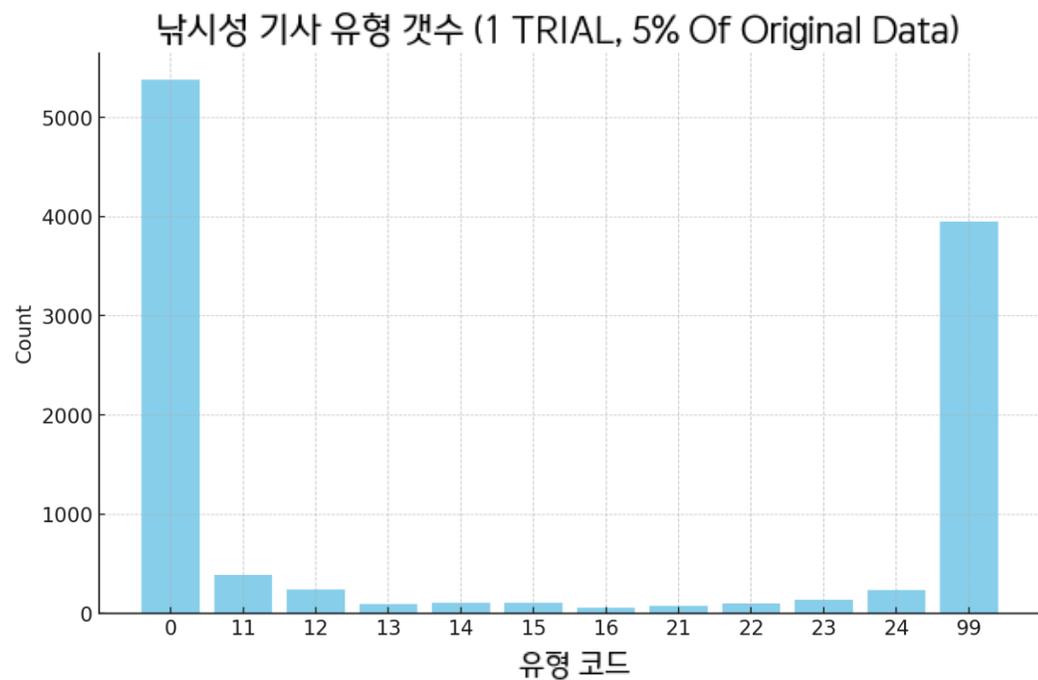
# EDA 및 전처리

낚시성 기사 유형 갯수



# 모델 훈련 과정

1 TRIAL	LLaMa2 7B Polyglot 12.8B	학습 데이터 크기 10,853 rows(기존 데이터 5%)	Polyglot 모델 선정
2 TRIAL	Polyglot ko 12.8B, 5.8B	학습 데이터 크기 12,000 rows (각 유형별로 1,000개씩)	학습 데이터 조정



# 모델 훈련 과정

3 TRIAL	Polyglot ko 5.8B, 1.3B	학습 데이터 크기 12,000 rows	학습률 조정 (모델 사이즈 별 테스트)
4 TRIAL	Polyglot ko 1.3B	학습 데이터 크기 12,000 rows	Parameter 조정

```
2 v training_args = TrainingArguments(  
3     output_dir = "outputs",  
4     per_device_train_batch_size=5,  
5     per_device_eval_batch_size=5,  
6     gradient_accumulation_steps=2,  
7     max_steps=1000,  
8     learning_rate=5e-3,  
9     fp16=True,  
10    logging_steps=10,  
11    optim="adamw_torch",  
12    report_to="tensorboard",  
13    lr_scheduler_type="cosine",  
14    warmup_steps=50,  
15    weight_decay=0.01,  
16    max_grad_norm=1.0,  
17    evaluation_strategy = "steps",  
18    eval_steps = 100,  
19    load_best_model_at_end = True,  
20    metric_for_best_model = "eval_loss" )
```

```
2 v training_args = TrainingArguments(  
3     output_dir = "outputs",  
4     per_device_train_batch_size=5,  
5     per_device_eval_batch_size=5,  
6     gradient_accumulation_steps=2,  
7     max_steps=-1,  
8     learning_rate=1e-5,  
9     fp16=True,  
10    logging_steps=10,  
11    optim="adamw_torch",  
12    report_to="tensorboard",  
13    lr_scheduler_type="cosine",  
14    warmup_steps=50,  
15    weight_decay=0.01,  
16    max_grad_norm=1.0,  
17    evaluation_strategy = "steps",  
18    eval_steps = 100,  
19    load_best_model_at_end = True,  
20    metric_for_best_model = "eval_loss" )
```

수행할 총 학습 단계 수  
(max\_steps),  
학습률 (learning\_rate)  
조정

# 모델 훈련 과정

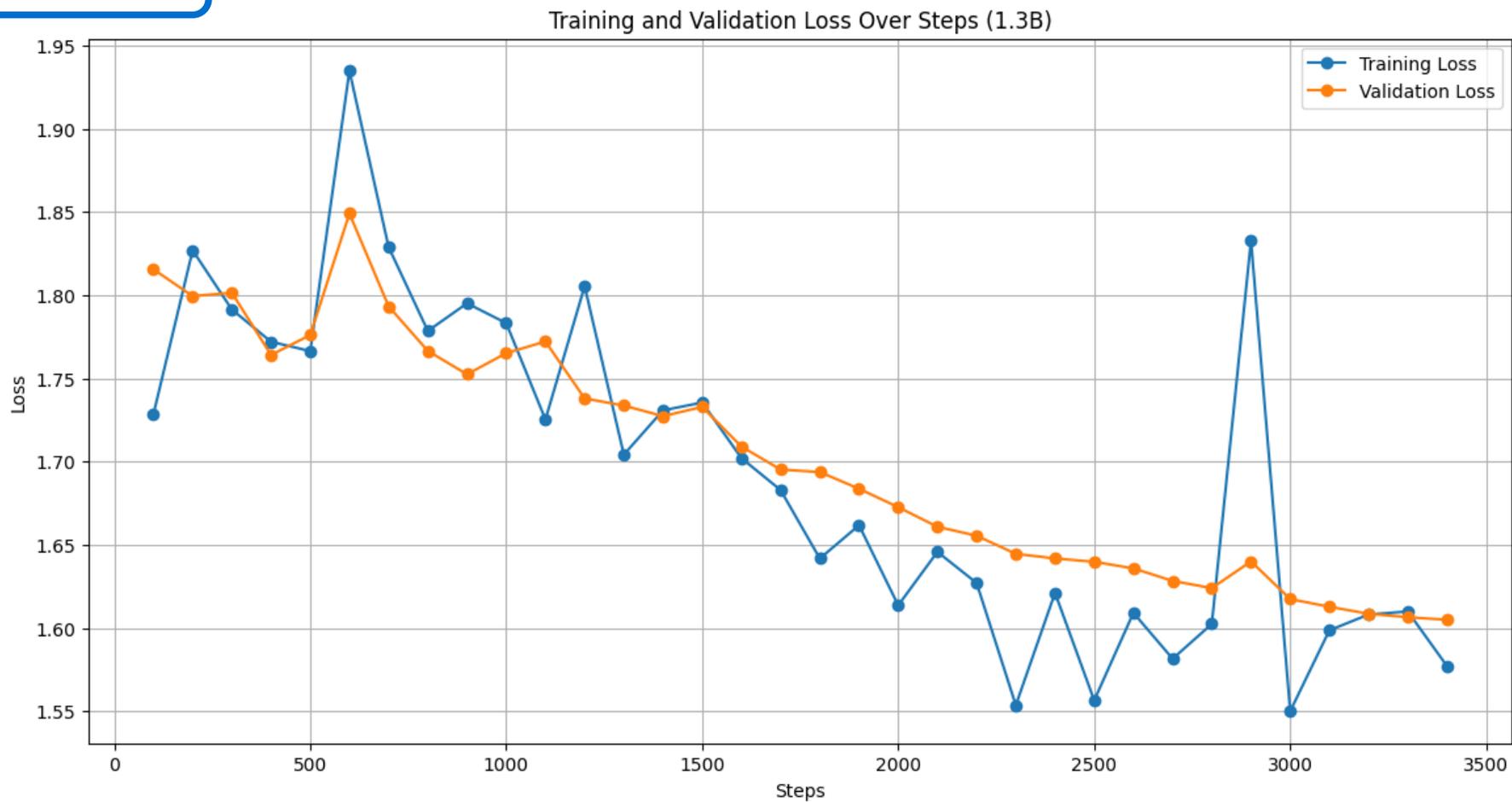
5 TRIAL	LLaMa2 7B (beomi) Polyglot ko 1.3B	학습 데이터 크기 12,000 rows	프롬프트 수정 Input text 특수문자 제거
---------	---------------------------------------	--------------------------	-------------------------------

## FINAL DATASET | DATA-CENTRIC APPROACH

Instruction Extension + $\alpha$	Pre-Instruction (string)	판별은 입력의 기사제목과 기사내용을 분석하여 해당 기사의 낚시성 기사 또는 정상기사, 낚시기사 유형을 출력합니다. 다음은 기사 제목, 기사 내용을 제공하는 입력과 짝을 이루는 판별 작업을 명령하는 지침입니다. 요청을 적절하게 완료하는 응답을 작성합니다.
	Instruction (string)	주어진 기사를 읽고 낚시성 기사 유무를 판별하라
Previous dataset	Input (string)	<p><b>기사제목:</b> f₩"내 통장에 웬 15억₩"...오송금 된 돈으로 호화생활 30대男의 최후</p> <p><b>기사내용:</b> f은행 실수로 하루아침에 백만장자가 된 남성이 징역 6년을 선고받아 억울함을 호소했다. 최근 뉴스위크는 지난해 6월 러시아 툴라에 있는 ATM에서 돈을 인출하는 과정에서 통장에 9520만 루블(약 14억 6700만원)을 발견한 남성의 사연을 보도했다. ... (후략)</p>
	Output (string)	해당기사는 낚시성기사입니다. 낚시기사 유형은 의문 유발형(부호) 입니다.

# 최종 학습 모델

POLYGLOT KO 1.3B



## 기사 요약

### 요약 정보

강하늘과 정소민이 로맨스 코미디가 아닌 코미디 로맨스로 강렬하게 돌아왔다. 30일은 드디어 D30, 서로의 찌질함과 뜯기를 견디다 못해 마침내 완벽하게 남남이 되기 직전 동반기억상실증에 걸려버린 정열(강하늘 분)과 나라(정소민)의 이야기를 담는다. 두 사람의 로맨스를 기대하고 온 이들은 당황할 수도 있다.

## 빈출 단어

### 빈출 빈도 수

word count	
0 웃음	12
1 코미디	6
2 영화	6
3 서로	6
4 정열	5



연예

## 김해준, 박세리와 ♥열애설 언급 "이성적으로 만난 것" 공개 발언까지

기사입력 2023.06.15 09:22 / 기사수정 2023.06.25 16:56

최희재 기자

공유



### 실시간 인기 기사

더보기

- | 연예                                | 스포츠 | 게임 |
|-----------------------------------|-----|----|
| 1 남현희, 돌연 4개의 입장문 "전청조, 계속 ...    |     |    |
| 2 '싱어게인3' 50호 김승미 빛투→반박, 진실...    |     |    |
| 3 임영웅, 드디어 '주제 파악' 완료..."눈물 날 ... |     |    |
| 4 남현희의 6개 입장문...대질신문 직전 '전청...    |     |    |

# 03. 결론 및 활용방안

## 서비스.

낱시성 기사 **유형 분류, 기사 요약, 빈출 단어 서비스**  
▶ 유저 친화적인 **UI 배포**

## 연구.

Model-centric & Data-centric  
▶ LLM Fine-Tuning을 위한 학습데이터 text-prompt  
**최적화 및 모델 학습 성능 향상**

# PRESSPULSE 확장성 제안

1. 단어 빈출 빈도로 감정 분석 후 **긍정/부정 기사 분류**
2. **댓글/조회수 등 추가 분석**으로 고도화된 판별 서비스 제공
3. 뉴스 포털 내 낚시성 기사 빈도 수 집계를 통한 **언론 지수 제시**
4. 낚시성 기사 필터링 후, **양질의 기사**를 제공하는 **플랫폼 서비스**

Q&A

감사합니다.