

Project Brainstorming Result

Project Name	한국어 기반 인공지능 생성 텍스트 감지
-----------------	-----------------------

17 조

202202501 조은비

202002494 박범창

202002565 좌진우

지도교수: 이종률 교수님 (서명)

# Document Revision History

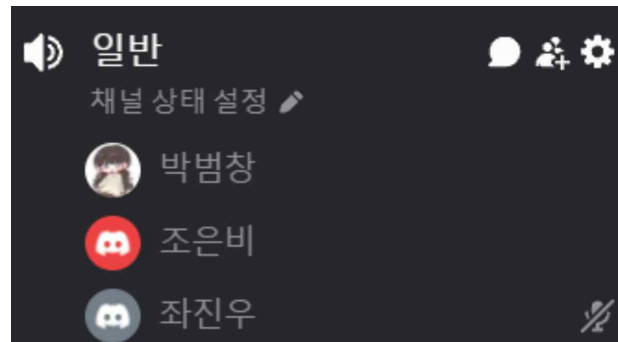
REV#	DATE	AFFECTED SECTION	AUTHOR
1	2025/03/28	아이디어 발산 및 수렴 작성	좌진우
2	2025/03/28	시각화하기 작성	박범창

# Contents

---

1. (문제 해결을 위한) 아이디어 발산.....	4
2. 아이디어 수렴.....	5
3. 시각화하기.....	6

## 1. (문제 해결을 위한) 아이디어 발산



## 2. 아이디어 수렴

핵심 개념	정의 및 하위 개념 목록
한국어 자연어 처리	<ol style="list-style-type: none"> <li>1. 허위 정보를 포함한 뉴스 기사를 자동으로 판별하는 기술</li> <li>2. 한국어 기반 모델 개발</li> <li>3. 형태소 분석, 문장 구조 파악 등 포함</li> </ol>
AI 생성 텍스트 탐지	<ol style="list-style-type: none"> <li>1. AI가 생성한 텍스트를 판별하는 기술</li> <li>2. 자연어 처리(NLP)와 머신러닝을 활용 또는 <b>fine-tuning</b> 사용</li> </ol>
탐지 모델 성능 평가	<ol style="list-style-type: none"> <li>1. 탐지 모델의 정확도, 재현율, 정밀도 등의 지표 활용</li> <li>2. 다른 기준 모델과의 성능 비교 분석</li> </ol>
가짜 뉴스 탐지	<ol style="list-style-type: none"> <li>1. 가짜 뉴스의 판별 기준 정의</li> <li>2. 각 도메인 별 신뢰도 평가 기준 설정</li> </ol>
탐지 회피 기법 대응	<ol style="list-style-type: none"> <li>1. 동의어 및 문장 구조 변환, 개인적인 경험 추가 등의 탐지 회피 기법에 대응 방안 마련</li> </ol>
데이터셋 구축	<ol style="list-style-type: none"> <li>1. 탐지 모델을 위한 대규모 한국어 데이터셋 확보</li> <li>2. AI Hub, 팩트체크 서비스 등의 자료 확보</li> </ol>

### 3. 시각화하기

