

# System Model (Sequence Diagram) Document

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

17 조

202202501 조은비

202002494 박범창

202002565 좌진우

지도교수: 이종률 교수님 (서명)

# Document Revision History

---

REV#	DATE	AFFECTED SECTION	AUTHOR
1	2025/04/30	Introduction, Usecase diagram 작성	좌진우

# Table of Contents

---

1.	INTRODUCTION .....	4
1.1.	연구 배경 .....	4
1.2.	연구 목적 .....	4
1.3.	연구 질문/ 가설 .....	4
2.	USE CASE DIAGRAM.....	5
2.1.	소프트웨어 활용 사례 .....	5
2.2.	문제 해결에 대한 사용 사례 DIAGRAM.....	6
3.	SEQUENCE DIAGRAM.....	7
3.1.	해결 방법에 대한 알고리즘 순서도 .....	7
4.	AI 도구 활용 정보 .....	9

# Introduction

---

## 1.1. 연구 배경

최근 가짜뉴스로 인한 피해사례가 급증하고 있어 전 세계적으로 가짜뉴스에 대한 대비책, 탐지방법 등이 연구되고 있음. 그러나 위와같은 연구는 대부분 영어권에 치중되어 한국어 기반에 알맞은 모델이나 연구방법이 아직 제대로 제시되지 못한 한계를 지니고 있음. 일례로 기존 연구 모델은 영어권에서는 90%이상의 정확도를 보여주지만, 이를 한국어에 그대로 적용할 경우 약 50~60%정도의 정확도를 보여 실생활에 적용하기엔 어려움이 존재함

## 1.2. 연구 목적

본 연구의 목적은 한국어기반 ai 생성 가짜뉴스 탐지를 통해 가짜뉴스를 정확히 탐지하는 방법을 연구하는 것이다 특히, 독자들이 자극적인 가짜뉴스에 휘둘려 피해를 보는 일을 해결하고 기사들의 신뢰도를 높이는 것을 목표로 한다. 이를 통해 가짜뉴스로 인한 피해가 줄고 건강한 정보사회를 만드는 것을 궁극적인 목표로 한다.

## 1.3. 연구 질문/ 가설

본 연구는 다음과 같은 연구 질문에 답하고자 한다:

- RQ1.

연구한 모델이 기존 연구모델에 비해 더 높은 정확도로 가짜뉴스를 판별하는가?

본 연구는 다음과 같은 가설을 가정한다

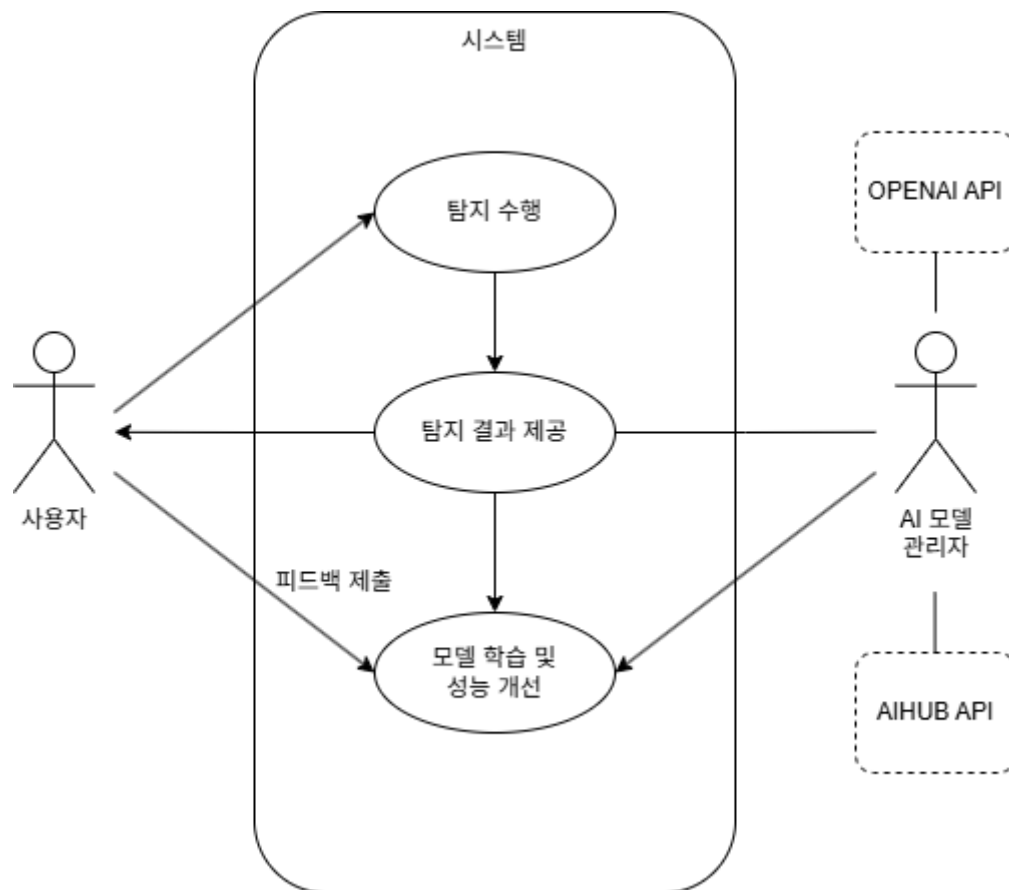
- H1.

AI 생성 기사를 만들기 위한 기초 기사 데이터는 AI로 생성된 기사가 아니다

-

## 2. Use Case Diagram

### 2.1. 소프트웨어 활용 사례



## 2.2. 문제 해결에 대한 사용 사례 Diagram

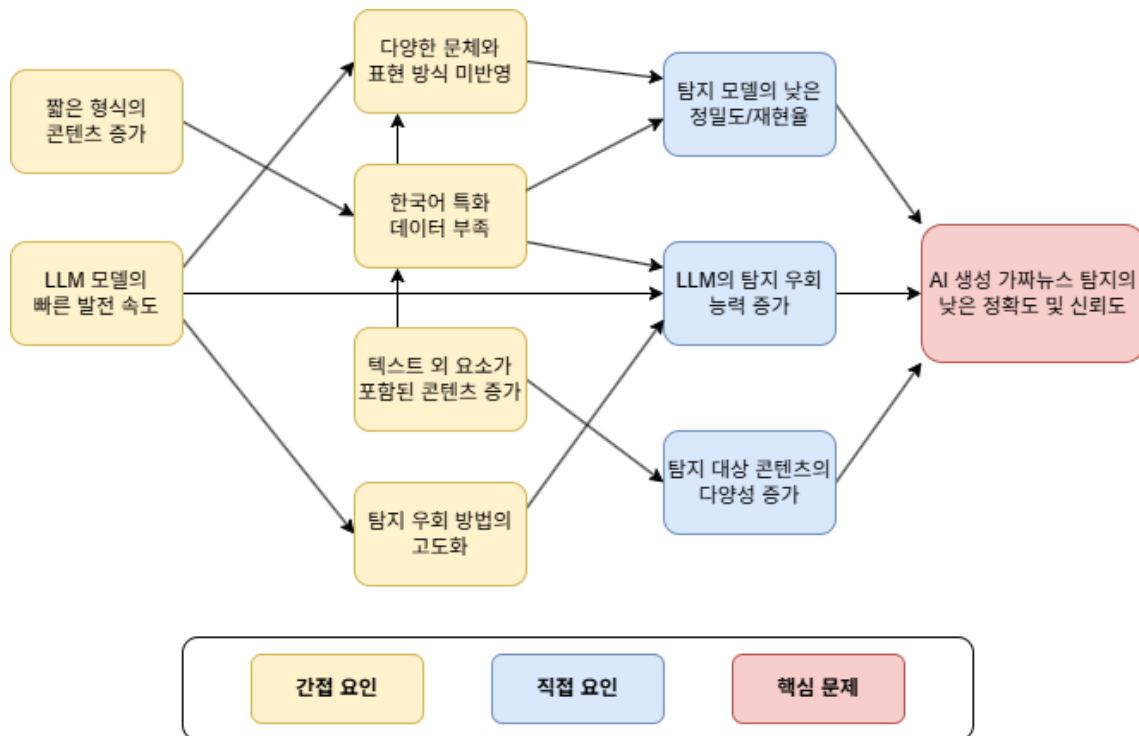
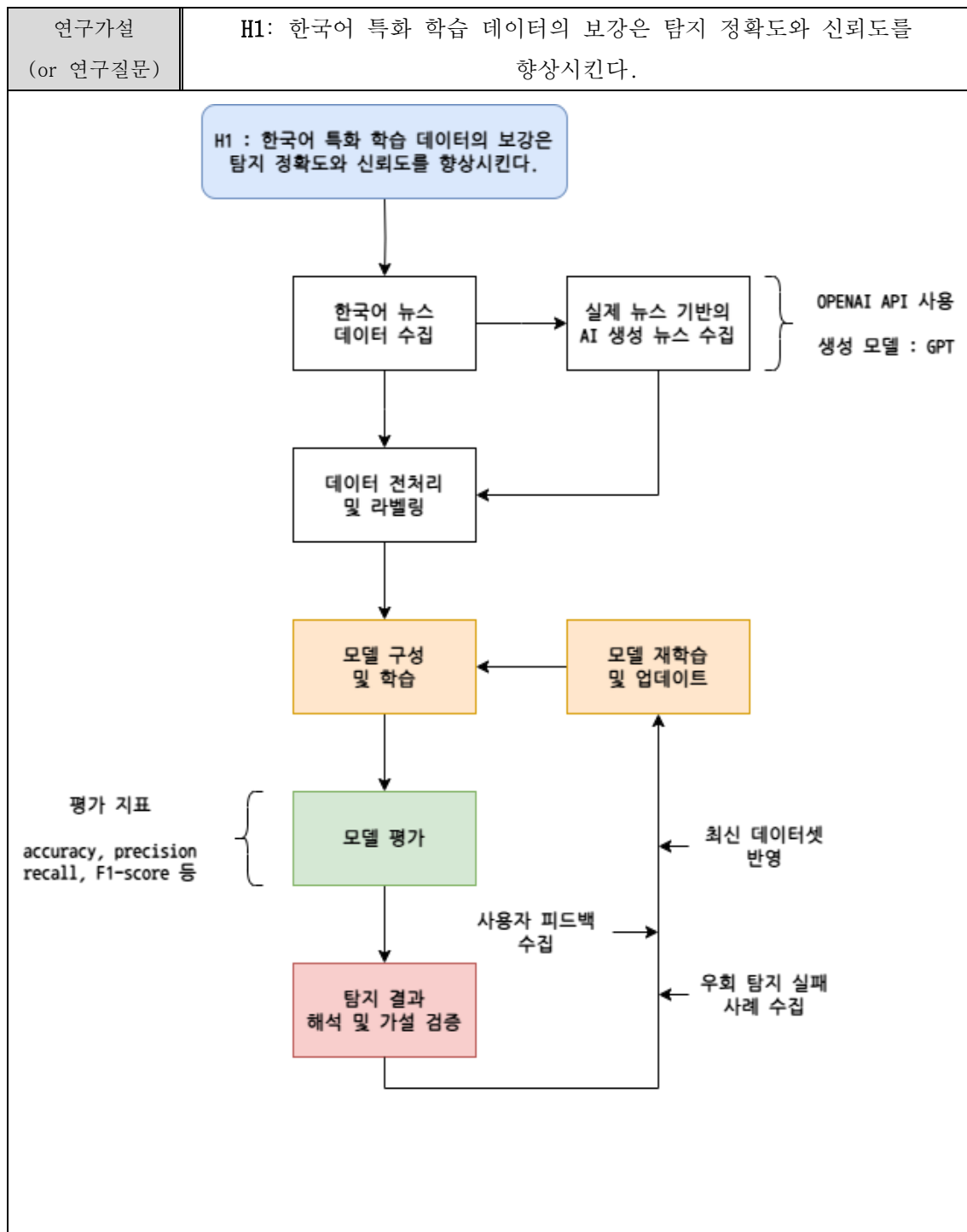


그림 1. 유스케이스 다이어그램

# Sequence Diagram

## 2.3. 해결 방법에 대한 알고리즘 순서도



핵심 문제 정의	AI 생성 가짜뉴스 탐지의 낮은 정확도와 신뢰도
알고리즘 순서	<ol style="list-style-type: none"> <li>1. 데이터 수집 및 AI 사용 뉴스 생성 <ul style="list-style-type: none"> <li>- AI Hub 뉴스 데이터 사용</li> <li>- AI 뉴스 생성을 위해 OPENAI API 사용</li> <li>- 생성 모델 : GPT</li> </ul> </li> <li>2. 데이터 전처리 및 라벨링 <ul style="list-style-type: none"> <li>- 진짜 / 가짜 뉴스 라벨링</li> <li>- 내용 생성이 안되거나 빈약한 데이터 삭제</li> </ul> </li> <li>3. 모델 구성 및 학습 <ul style="list-style-type: none"> <li>- KoBART, KoELECTRA 모델을 사용</li> </ul> </li> <li>4. 모델 평가 <ul style="list-style-type: none"> <li>- 정확도, 정밀도, 재현율, F1-score 등의 지표 활용</li> <li>- KoELECTRA, KoBART 등의 모델마다 각 지표 측정 및 평가</li> </ul> </li> <li>5. 탐지 결과 해석 및 가설 검증 <ul style="list-style-type: none"> <li>- 탐지 결과를 해석하고, 해당 결과가 가설을 뒷받침하는지를 검증</li> <li>- 사용자 피드백, 우회 탐지 실패 사례를 수집하고, 최신 데이터셋을 반영하여 발전하는 LLM의 우회 능력에 대응할 수 있도록 함</li> </ul> </li> <li>6. 모델 재학습 및 업데이트 <ul style="list-style-type: none"> <li>- 빠르게 발전하는 인공지능에 대항할 수 있도록 지속적인 학습과 피드백 적용</li> </ul> </li> </ol>



### 3.AI 도구 활용 정보

사용 도구	GPT-4, Claude 2.1
사용 목적	인터뷰 질문 초안 작성, 문장 흐름 정리, 사례 리서치 보조
프롬프트	<ul style="list-style-type: none"> <li>● 사용자 경험 중심으로 문제정의서 예시를 보여줘</li> <li>● 디자인 브레인스토밍 아이디어 10개 제안해줘</li> </ul>
반영 위치	<ol style="list-style-type: none"> <li>1. 인터뷰 질문 목록 (p.5)</li> <li>2. 아이디어 설명 문단 정리 (p.6)</li> </ol>
수작업	있음(논리 보강, 사례 교체 등)
수정	