

코딩 가능 수준(or 청소년 레벨) 활동 가능 리스트

(2번은 완성됨, 다만 코딩가능 수준 조 월요일에 확인후, 2개 제공할지 3개 제공할지 결정)

## 1. 간단한 ChatGPT API 활용 챗봇 만들기

**활동:** chatgpt api + streamlit을 이용하여, 간단한 나만의 챗봇을 만들고

+prompt 를 학생들이 직접 넣게 하여, 커스텀화 하기

예를들어: 수학문제 가르쳐주는 교사 로봇, 변호사를 만들어 강아지 체포에 대한 변호하기..

**장점:** 구현 난이도에 비하여 다양한 기능을 체험할수 있음

**단점:** 머신러닝, deep learning 보단, generative Ai 에 집중되어 있음.

Later..

## 2. 간단한 Sentiment Analysis(감정분석)

### 2.1 스팸메일 구분, 2.2 긍정/부정 구분 (영화 리뷰 or sns 문자)

- 개요

- SNS 댓글 데이터나 간단한 스팸/햄 이메일 데이터셋 등을 사용, 긍정/부정(이진 분류) 모델을 학습.
- Scikit-learn의 CountVectorizer, TfidfVectorizer로 텍스트를 간단히 벡터화하여 Logistic Regression, Naive Bayes 등으로 빠르게 학습 가능. (해당 라이브러리 설명은 진행하지 않음)

- 액티비티 계획

- 이미 정리된 간단한 CSV 데이터셋 사용 or 커스텀 데이터 list 미리 제공 후
- 학생들이 추가로 몇개를 넣으면서 연습하게 하기 or spam/ham 인지 직접 데이터 라벨 붙이게하고,
- 일부 코드 완성시키기 (optional)
- (“스팸/햄” 레이블이 있는 SMS 예제 등).
- ~~벡터화 + 모델 학습 + 예측 결과 혼동 행렬(Confusion Matrix) 정도 간단히 시각화.~~(제공됨)
- 추가 고려 사항 (부모님 발표를 위하여 text to speech로 음성 인식 시키기 or streamlit등을 이용하여.

간단하게 터미널이 아닌 UI화면 제공하기

- 발표 포인트

1. 이 문장이 스팸인지 아닌지 구별한다

2. “부모님이 입력한 문장이 긍정/부정인지 AI가 맞춰본다.”

### 3. 간단한 음성 인식 or 음성 분류 (시간이 매우 촉박하다면 시연 위주)

- 개요
  - Google Speech-to-Text API 같은 것을 활용해서 음성을 텍스트로 변환 → 그 텍스트에 대해 추가 분석(감정분석 등).
  - 또는 음성의 특성(스펙트로그램)으로 간단한 분류(예: 음악 장르 분류)를 시연.
- 장점
  - “말을 하면 AI가 이해한다”라는 시연 자체가 임팩트 있음.
- 주의점
  - 음성 인식은 외부 API 사용하면 간단하지만, 직접 딥러닝 모델을 학습하려면 시간이 많이 걸림(데이터셋, GPU 등).
- 발표 포인트
  - “이게 머신러닝(딥러닝)이 실제 음성을 어떻게 처리해 텍스트 또는 카테고리를 추출하는지” 간단한 원리 소개