

PPT 기반 학습 보조 시스템: 파싱·키워드 추출·GPT 기반 과목 유추 및 수준별 개념 설명

윤영웅, 윤태현, 양효진, 윤숙
국립목포대학교 컴퓨터공학과
e-mail : hyojin980417@gmail.com

PPT-Based Learning Support System: Parsing, Keyword Extraction, GPT-based Subject Inference, and Level-Specific Concept Explanations

Yeong woong Yoon, Tae hyeon Yoon, Hyo Jin Yang, Yoon sook

Computer Engineering, Mokpo National University

요 약

본 연구는 대학생 학습을 보조하기 위한 PPT 기반 학습 지원 시스템을 제안한다. 본 시스템은 강의 자료(PPT)를 업로드하면 자동으로 슬라이드 텍스트를 파싱하고 핵심 키워드를 추출한다. 이어 GPT 기반 과목 유추 기능을 통해 PPT가 어떤 전공 과목에 속하는지 판별하며, 사용자는 직접 수준(하/중/상)을 선택하거나 테스트 모드를 수행하여 학습 수준을 진단하여 최종적으로 유추된 과목과 수준에 따라 각 슬라이드별 핵심 단어를 강조하여 개념을 설명하며, 결과는 ZIP 파일로 다운로드할 수 있다.

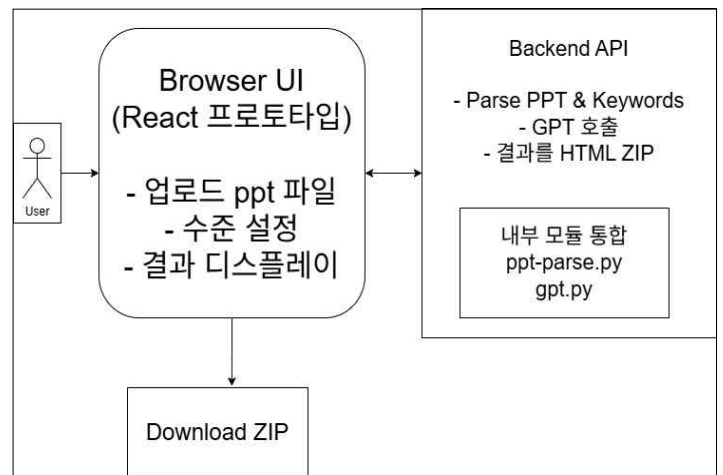
1. 서 론

대학 교육 환경에서 PPT 강의 자료가 주요 학습 수단으로 활용되면서, 방대한 양의 슬라이드를 효과적으로 복습하고 핵심 개념을 파악하는 데 어려움이 존재한다. 본 연구는 이러한 문제점을 해결하기 위해 PPT 기반 학습 지원 시스템을 설계 및 구현하였다. 제안하는 시스템은 PPT 파일을 업로드하면 텍스트 파싱, 키워드 추출, GPT 기반 과목 유추를 수행한다. 이후 사용자 수준에 맞는 개념 설명을 제공하여 개인 맞춤형 학습을 지원하며, 결과물은 ZIP 파일로 다운로드할 수 있다.

2. 관련 연구

2.1 시스템 아키텍처

본 시스템은 백엔드(Flask)와 프론트엔드(React)로 구성된 클라이언트-서버 아키텍처를 채택하였다. PPT 파싱, 키워드 추출, GPT 통신 등은 각각의 모듈로 분리하여 응집도를 높였다. 시스템 전체 구성도는 그림 1과 같다.



(그림 1) 시스템 구성도

2.2 주요 기능

본 시스템의 주요 기능은 사용자의 학습 흐름에 따라 순차적으로 동작하며, 총 8단계로 구성된다.

첫째, 'PPT 파일 파싱' 기능은 사용자가 업로드한 PPT 파일(.pptx)을 대상으로 python-pptx 라이브러리를 이용해 슬라이드별 텍스트 데이터를 추출한다. 이 과정에서 불필요한 공백이나 개행문자를 정제하여 텍스트 구조를 분석한다.

둘째, '핵심 단어 추출' 기능은 YAKE 알고리즘을 활용하여 파싱된 텍스트 내에서 주요 단어를 자동으로 추출하며, 이를 통해 슬라이

드별 핵심 키워드 리스트를 생성한다.

셋째, '과목 유추' 기능은 추출된 핵심 키워드를 기반으로 GPT 프롬프트를 구성하여 API에 전달하며, 해당 PPT가 어떤 전공 과목에 속하는지 자동으로 판별한다.

과목이 판별된 후, 네 번째 단계인 '사용자 학습 모드 선택'을 통해 사용자는 학습 난이도를 설정한다. 사용자는 A 모드(수준 직접 선택) 또는 B 모드(테스트) 중 하나를 선택할 수 있다.

다섯째, B 모드(테스트)를 선택한 경우 '수준 평가' 기능이 활성화된다. 이는 GPT 기반의 자동 평가 방식으로, GPT가 동적으로 생성한 객관식 5문항을 통해 학습자의 이해 수준을 '상·중·하'의 세 단계로 분류한다.

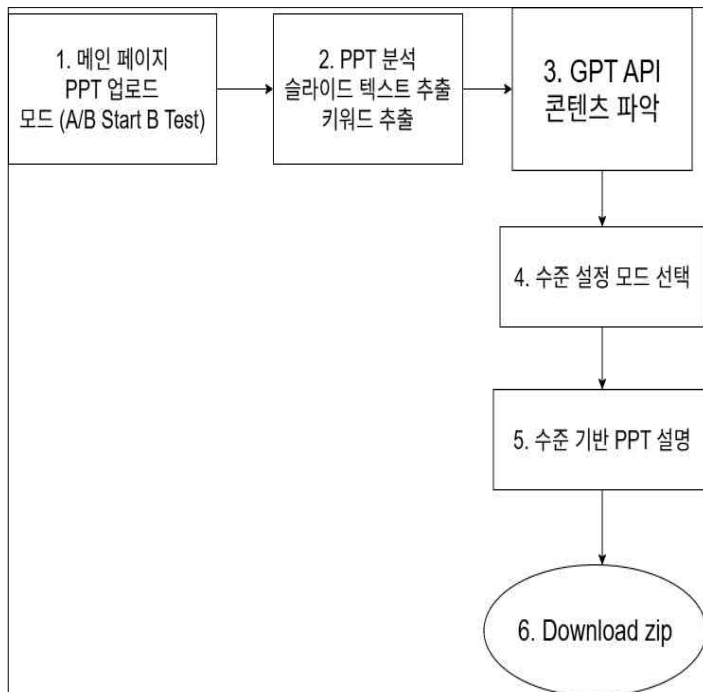
여섯째, 학습 수준이 확정되면 시스템의 핵심 기능인 '개념 설명 생성'이 이루어진다. 이는 앞서 유추된 과목명과 확정된 사용자 수준을 기반으로, GPT가 각 핵심 단어를 중심으로 맞춤형 개념 설명을 생성하는 기능이다.

일곱째, '결과 시각화' 기능은 생성된 모든 데이터를 사용자에게 직관적으로 제공한다. 웹 화면 좌측에는 원본 슬라이드 내용을, 우측에는 해당 슬라이드의 핵심 단어와 GPT가 생성한 맞춤형 개념 설명을 함께 표시한다.

마지막으로, '결과 다운로드' 기능은 사용자가 결과 페이지의 내용을 오프라인에서도 확인할 수 있도록, HTML 및 JSON 파일이 포함된 ZIP 형식으로 다운로드할 수 있게 구현되었다.

참고문헌

- [1] GPT-4o Model (2024), OpenAI, (accessed Sep., 20, 2025).
- [2] Python-pptx: A Python library for creating and modifying PowerPoint files, <https://python-pptx.readthedocs.io/> (accessed Sep., 12, 2025).
- [3] YAKE: Yet Another Keyword Extractor, <https://github.com/LIAAD/yake> (accessed Sep., 14, 2025).
- [4] "대학생의 생성형 인공지능 활용에 관한 연구 : ChatGPT를 중심으로". 한국교육정보미디어학회, 2024
- [5] "개별화 맞춤형 학습을 지원하는 AI 챗봇 기반 플랫폼 분석". 디지털콘텐츠학회논문지, 2024.
- [6] E-QGen: Educational Lecture Abstract-based Question Generation System (강의 초록 기반 질문 생성 시스템). themoonlight.io (논문 리뷰).
- [7] "영어학습 플랫폼으로서의 ChatGPT 활용 방안". 멀티미디어 언어교육, 2023.
- [8] "Automatic question generation: a review of methodologies, datasets, evaluation metrics, and applications (자동 질문 생성 리뷰)". Journal of Big Data (PMC), 2023.
- [9] "Chat GPT and education (Chat GPT와 교육)". ResearchGate, 2023.
- [10] "Utilizing GPT for Interactive Dialogue-based Learning Scenarios (대화형 학습 시나리오를 위한 GPT 활용)". DiVA (KTH), 2024.



(그림 2) 시스템 흐름도