

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

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

윤영웅, 윤태현, 양효진, 윤숙

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

요약

본 연구는 대학생 학습을 보조하기 위한 PPT 기반 학습 지원 시스템을 제안한다. 본 시스템은 강의 자료(PPT)를 업로드 하면 자동으로 슬라이드 텍스트를 파싱하고 핵심 키워드를 추출한다. 이어 GPT 기반 과목 유추 기능을 통해 PPT가 어떤 전공 과목에 속하는지 판별하며, 사용자가 직접 수준(하/중/상)을 선택하거나 자동 테스트 모드(객관식 퀴즈 5문항)를 수행하여 학습 수준을 진단한다. 최종적으로 유추된 과목과 수준에 따라 각 슬라이드별 핵심 단어를 강조하여 개념을 설명하며, 결과는 ZIP 파일로 다운로드할 수 있다. 본 연구는 팀 프로젝트 형태로 구현되었으며, 백엔드와 프론트엔드 간 연동 및 사용자 친화적 학습 지원을 목표로 한다.

■ 중심어 : PPT 파싱, 키워드 추출, 과목 유추, GPT, 학습 수준 평가

Abstract

This study proposes a PPT-based learning support system to assist university students. When a user uploads lecture materials (PPT), the system automatically parses the slide text and extracts key keywords. Through a GPT-based subject inference function, it then determines the academic subject of the presentation. Users can diagnose their learning level by either directly selecting a level (beginner, intermediate, advanced) or by completing an automatic test mode consisting of five multiple-choice questions. Finally, based on the inferred subject and the user's level, the system provides concept explanations for each slide, highlighting the key words. The results can be downloaded as a ZIP file. This system was implemented as a team project, aiming for seamless integration between the backend and frontend to provide user-friendly learning support.

■ keywords : PPT Parsing ; Keyword Extraction, Subject Inference, GPT, Learning Level Assessment

I. 서론

오늘날 대학 교육 환경에서는 학습 자료의 디지털화가 보편화되면서, 특히 PowerPoint(PPT) 형태의 강의 자료가 주요 학습 수단으로 활용되고 있다. 하지만 방대한 양의 PPT 슬라이드를 효과적으로 복습하고 핵심 개념을 파악하는 데 어려움을 겪는 학생들이 많다. 또한, 학업 성취도에 따라 맞춤형 학습 자료를 제공받기 어려운 현실적인 문제도 존재한다. 본 연구는 이러한 문제점을 해결하기 위해 PPT 기반 학습 지원 시스템을 설계하고 구현하였다. 제안하는 시스템은 사용자가 PPT 파일을 업로드하면 텍스트를 자동으로 파싱하고 핵심 키워드를 추출한다. 또한, 추출된 키워드를 바탕으로 GPT 모델을 활용하여 해당 PPT가 어떤 전공 과목에 속하는지 유추한다. 사용자의 학습 수준을 진단한 후, 유추

된 과목과 수준에 맞는 개념 설명을 제공하여 개인 맞춤형 학습을 지원한다. 최종 결과물은 웹 UI에서 시각적으로 확인할 수 있으며, ZIP 파일 형태로 다운로드할 수 있도록 하여 사용 편의성을 높였다.

II. 본론

1. 시스템 설계 및 구현

본 시스템은 PPT 파싱, 키워드 추출, GPT 기반 과목 유추, 학습 수준 평가, 개념 설명, 그리고 결과 출력의 여섯 가지 주요 기능을 포함한다. 시스템은 백엔드(Flask)와 프론트엔드(HTML/JS, React)로 구성된 클라이언트-서버 아키텍처를 채택하였다.

2. 주요 기능 및 구현

F-1. PPT 업로드 및 파싱: 사용자가 업로드한 PPT 파일

(.pptx)의 슬라이드 텍스트를 Python의 python-pptx 라이브러리를 이용하여 효율적으로 추출한다.

F-2. 키워드 추출: 추출된 텍스트에서 논문의 주요 내용을 반영하는 핵심 키워드를 YAKE 또는 TextRank와 같은 키워드 추출 알고리즘을 활용하여 자동화한다.

F-3. 과목 유추: 추출된 키워드와 슬라이드 내용을 기반으로 GPT-40 모델에 프롬프트를 전달하여 해당 PPT가 어떤 전공 과목에 속하는지 판별한다.

F-4. 사용자 수준 평가: 사용자의 학습 수준을 파악하기 위해 두 가지 모드를 제공한다. 모드 A: 사용자가 직접 학습 수준(하/중/상)을 선택한다. 모드 B: GPT가 생성한 5개의 객관식 퀴즈를 풀고, 그 결과를 바탕으로 학습 수준을 자동 등급 산출한다.

F-5. 개념 설명: 유추된 과목과 평가된 수준에 따라 각 슬라이드별 핵심 단어에 대한 개념 설명을 GPT로부터 생성하여 제공한다.

F-6. 결과 출력: 생성된 학습 결과는 웹 UI에서 시각적으로 보여주며, 사용자가 필요시 모든 내용을 ZIP 파일로 다운로드할 수 있도록 한다.

3. 시스템 아키텍처

시스템은 백엔드를 위한 Flask REST API와 프론트엔드를 위한 React(초기 프로토타입은 HTML/JS)로 구성된다. PPT 파싱, 키워드 추출, GPT 통신 등은 각각 ppt_parser.py, keywords.py, gpt.py 모듈로 분리하여 역할의 응집도를 높였다.

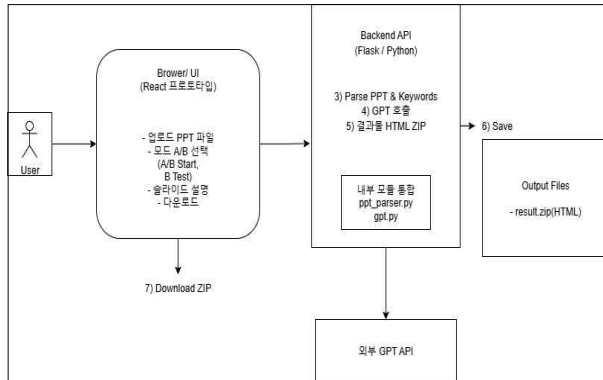


그림 1 PPT 기반 학습 지원 시스템 전체 구성도

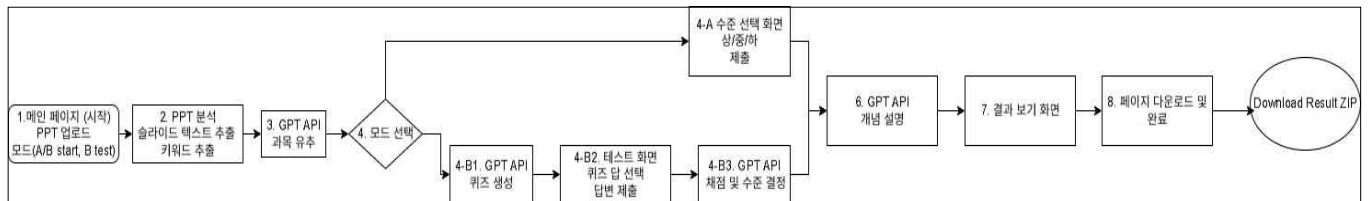


그림 2 시스템 처리 흐름도

III. 결론

본 연구는 GPT 기반 PPT 학습 지원 시스템을 설계 및 구현하여 대학생들의 효과적인 학습을 돕고자 하였다. 이 시스템은 기존의 학습 지원 도구가 제공하지 못했던 개인 맞춤형 개념 설명을 자동화함으로써, 사용자가 복잡한 강의 자료의 핵심을 더 쉽게 파악하고 능동적인 학습을 할 수 있도록 지원한다. 향후 연구에서는 더 다양한 파일 형식(PDF, DOCX 등)을 지원하고, 학습 피드백 시스템을 고도화하여 사용자 경험을 더욱 향상시킬 계획이다.



그림 3 학습 요약 결과 보고서

REFERENCES

- [1] GPT-40 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).