Slidify.ai: Leveraging AI for Sophisticated Presentation Drafts

Slidify.ai Encode AI Hackathon

March 10, 2024

Abstract

Slidify.ai is an innovative AI application designed to convert LaTeX research papers into compelling presentation drafts. Unlike existing solutions, Slidify.ai offers a nuanced and tailored approach by leveraging Large Language Models (LLMs) and generative AI techniques. This paper presents the methodology, technical stack, and results of Slidify.ai, along with future directions for development.

1 Background

Slidify.ai project is an AI application that is able to convert research papers in LaTex form to a presentation draft. This use case has been explored countless times by various companies, even those from Big Tech (OpenAI, Google Slides, Canva, etc.) but has often resulted in products that don't meet the needs of users due to overly simplistic or generalized outputs. Leveraging Large Language Models (LLMs) for extracting essential content like key information, points, tables, and figures from research papers, and combining this with generative AI for visual presentation, our application promises a more nuanced and tailored approach to this problem. By also incorporating user-driven customization options, such as the choice in the number of slides and preferred templates, slidify.ai not only enhances the usability of the generated presentations but also aligns closer with the specific requirements of the target audience. This thoughtful integration of AI capabilities and user customization stands to significantly bridge the existing gap in this domain.

2 Introduction

Due to the shortcomings of even big tech companies in this aspect, we decided to take on the challenge of creating a better version of this idea. For slidify.ai, we have created an application that allows users to upload Latex documents and convert them using AI to presentations. Our novel idea is that the application will extract the tables, figures, and information that people use for presentation, instead of providing generalized bullets with a generic slide.

3 Methodology

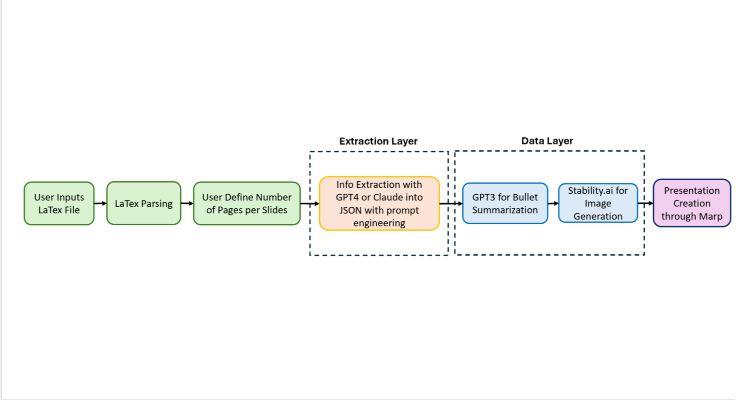

Extraction Layer
Data Layer
User Inputs LaTex File
LaTex Parsing
User Define Number of Pages per Slides
Info Extraction with GPT4 or Claude into JSON with prompt engineering
GPT3 for Bullet Summarization
Stability.ai for Image Generation
Presentation Creation through Marp

Figure 1: Flowchart of the Slidify.ai process

The methodology of Slidify.ai consists of four main stages: LaTeX parsing and information extraction, extraction layer, data layer, and presentation creation. Users upload LaTeX files, which are processed using LLMs to extract key information such as titles, section names, tables, and figures. Additional data for presentation, including bullet points and images, are generated using AI techniques. Finally, presentations are created using the Marp library.

3.1 LaTex Parsing and Information Extraction

In this process, users will first upload a LaTex File to the application. The LaTex file is then processed using a script to allow for easier extraction by the AI. Before extraction, users are given the option to select the number of pages they would like for each section.

3.2 Extraction Layer

Here, the extraction process begins. This process is carried out by an LLM, and we have programmed the application to work with both OpenAI's GPT4 or Anthropic's Opus. The LLMs are given a set of detailed and specific prompts based on the paper's information, giving it the ability to extract key information from the papers. This information includes the paper title, section names, section information, tables, and figures which will then be precisely entered into a specific JSON format by the LLM output. To achieve this, we utilised prompt techniques and successfully obtained results that were accurate and consistent. Additionally, speaker notes, or prepared speeches for each slide will also be provided at this stage.

3.3 Data Layer

In this section, additional data for presentation will be generated. Firstly, bullet points for the slides will be generated by summarising the speaker notes with GPT-3.5. Then, the application will generate images for sections that did not have figures in the paper. This will be done using Stability.ai's API.

3.4 Presentation Creation

Finally, presentations will be created in this layer. We opted to use Marp as our library as it worked better at creating great looking slides compared to many other recommended libraries. In the end, we provide users with the option to download our generated slides as pdf.

4 Technical Stack Summary

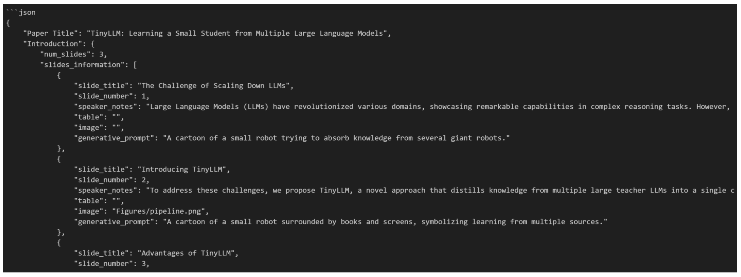
Figure 2: Technical Stack

Slidify.ai utilizes Streamlit for the front-end, Python for the backend (potentially utilizing libraries like PyLaTeX), Docker for deployment, and Google Cloud for hosting. The application integrates Stability.ai's API for image generation.

5 Results

The main contribution of our work was engineering the prompts for our LLM, which was able to extract all the information we needed in the paper and represent them as json output. This is not something that has been seen in other previous works. We present a figure of our results below, showing our json output from the extraction.

From the results, we found that OpenAI's GPT4 and Anthropic's Claude were able to extract a variety of Latex papers with perfect accuracy and consistency. This is a big achievement, as alternatives like regex do not have a generalizable pattern for all cases, and it will definitely not be able to pick up



~~~
```json
{
    "Paper Title": "TinyLLM: Learning a Small Student from Multiple Large Language Models",
    "Introduction": {
        "num_slides": 3,
        "slides_information": [
            {
                "slide_title": "The Challenge of Scaling Down LLMs",
                "slide_number": 1,
                "speaker_notes": "Large Language Models (LLMs) have revolutionized various domains, showcasing remarkable capabilities in complex reasoning tasks. However,
                "table": "",
                "image": "",
                "generative_prompt": "A cartoon of a small robot trying to absorb knowledge from several giant robots."
            },
            {
                "slide_title": "Introducing TinyLLM",
                "slide_number": 2,
                "speaker_notes": "To address these challenges, we propose TinyLLM, a novel approach that distills knowledge from multiple large teacher LLMs into a single c
                "table": "",
                "image": "Figures/pipeline.png",
                "generative_prompt": "A cartoon of a small robot surrounded by books and screens, symbolizing learning from multiple sources."
            },
            {
                "slide_title": "Advantages of TinyLLM",
                "slide_number": 3,
~~~

Figure 3: Json Results of LLM Output

semantic information in the paper. Semantic understanding is important especially during figure extractions, as figures sometimes are not placed in the same sections as where they are mentioned. Thus, our model being able to identify these nuances were also notable. .

6 Future Work

In the future, we propose to expand the idea that we propose to wider applications beyond Latex Files. They could include normal documents, essays, finance reports, etc. Besides that, we also believe that we can further improve our application by adding an image generation feature for the background and overall theme of the slides as well. Lastly, our most novel idea is to potentially create a 'Research Paper to Educational Shorts Video AI'. How this could be done is by first feeding the speaker notes we generated into a speech to text AI, then stitching the slides together to form a video. This could potentially revolutionise the entertainment, educational, and leisure industry.

7 Conclusion

Slidify.ai leverages NLP technology to convert complex academic papers into persuasive presentation slides, addressing the challenge faced by researchers and students in disseminating their work. By automating the slide creation process, Slidify.ai empowers users to focus more on their research activities while effectively sharing their findings with a broader audience.

Acknowledgment

We would like to thank Stability.ai for their support and providing the image generation API.