UNIVERSITY OF
TORONTO

PowerGenX:
Automated PowerPoint Creation By GPT-4

Alex Wang - alexei.wang@mail.utoronto.ca

Johnny Gan - johnny.gan@mail.utoronto.ca

Course Project

ECE1786H F

Creative Applications of Natural Language Processing

Word Count: 1999 words

Penalty: 0%

December 8, 2023

Contents

1 Introduction

PowerGenX is a project that simplifies the process of creating PowerPoint presentations structure compared with normal manually creating. Utilizing the capabilities of OpenAI's GPT-4, PowerGenX will transform standard text documents into comprehensive, visually appealing presentations. The project's primary objective is to create PowerPoint slides based on user preferences and documents while following the principles of concise and engaging presentation design.

Central to PowerGenX's functionality is its capacity to parse and utilize input data from either .docx or .pdf file formats. These inputs could range from professional proposals and reports to academic articles. Utilizing this data, PowerGenX crafts well-organized slide outlines by GPT-4 Interaction Module, it will also generate the presentation creation code in cooperation with the PowerPoint Generator Class that contains the clear functions to generate a PowerPoint. Additionally, PowerGenX integrates the Pexels API to source and embed relevant imagery within the slides. This feature significantly benefits users who need quick, yet impactful presentations, as it automates the often time-consuming task of manual image selection and integration.

Figure 1: Core Idea of PowerGenX

2 Data and Data Processing

In the PowerGenX project, the data refers to the textual input provided by the user, which is pivotal in generating tailored PowerPoint presentations. This input, typically a text document, is either in .docx or .pdf format. The document contains the intended content for the PowerPoint slides, ranging from proposals to educational material. Below is a screenshot of our frontend that collects user's input. At left side user choose to upload their documents and adding more notes, then some preference check boxes are provided for a personal design.

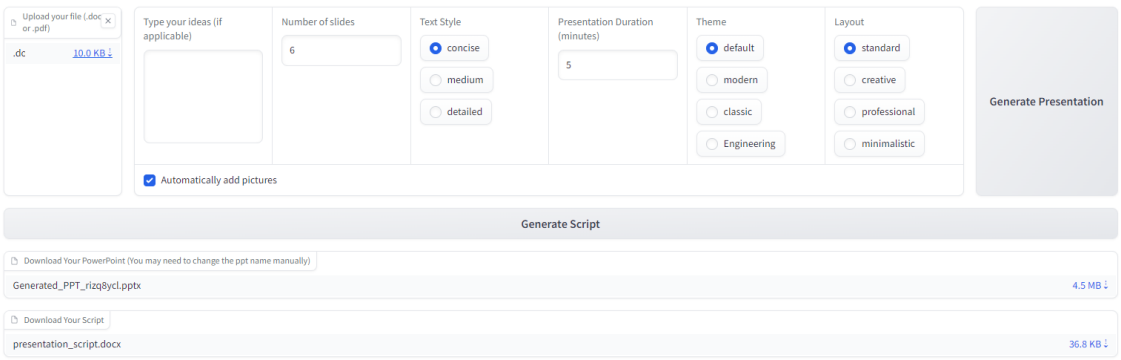

Figure 2: Gradio frontend of PowerGenX

2.1 Data Colleciton

The data collection process in PowerGenX is user-driven. Users upload their documents, which serve as the primary data source. For instance, a document titled "AI in Healthcare" might include sections such as an introduction to AI, its healthcare applications, challenges, and future prospects. Each section corresponds to specific aspects of the presentation, forming a structural backbone for slide generation.

2.2 Data Processing

The PowerGenX system processes this input data using the GPT-4. GPT-4 analyzes the text, understanding its context and structure, and generates a coherent outline for PowerPoint slides. This outline includes slide titles, bullet points, and detailed content that reflects the document's key themes.

3 Background & Related Work

Typeset is a notable example of a related work in the field of automated PowerPoint creation. [1] It is mainly done by inputting an idea or a document and then generating PowerPoint through AI. Which is similar to that, typeset can customize their PowerPoint according to users' different requirements. However, typeset has a huge database of templates, so it can generate rich backgrounds and pictures according to different needs. Here is the interface of the Typeset:

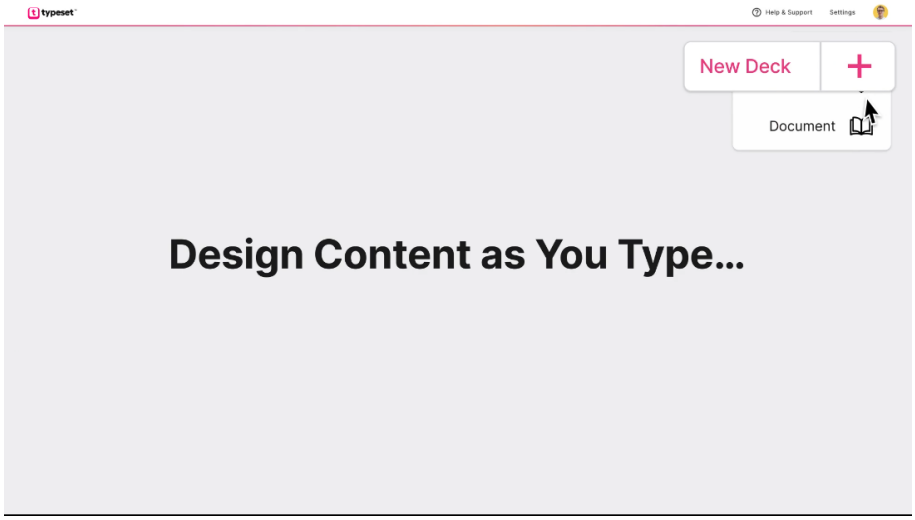

Figure 3: Typeset Interface

4 Architecture

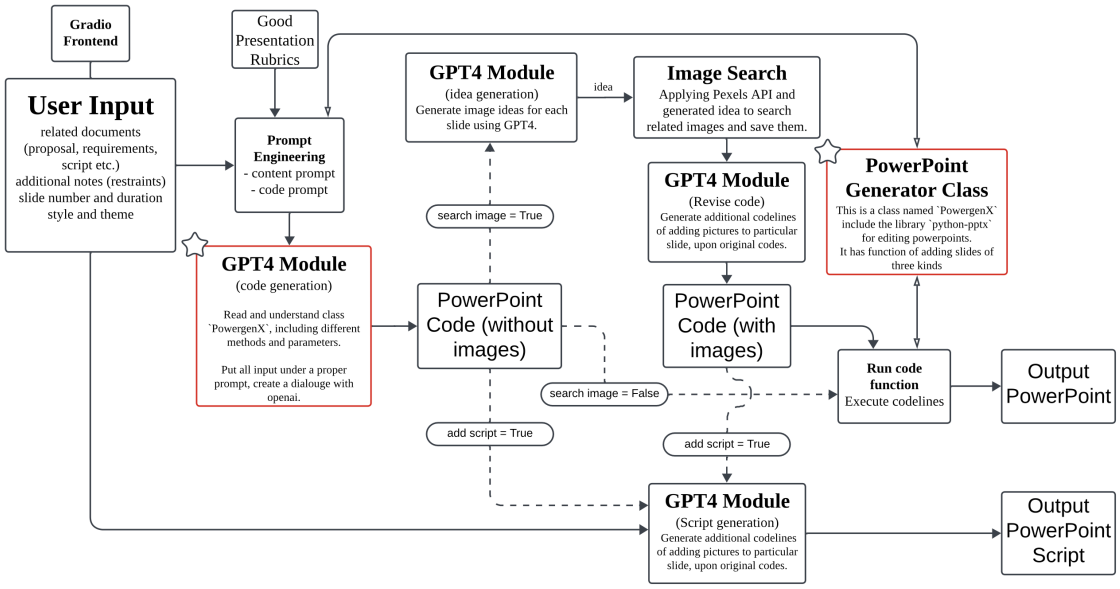

Figure 4: Main Model

4.1 User Input Acquisition

At the system's outset, the Gradio frontend interface captures the user's input, which encompasses the primary content for the slides in the form of text documents and supplemental specifications such as presentation style, theme, and desired number of slides.

4.2 GPT-4 Interaction & Prompt Engineering

Prompt engineering is an essential aspect of working with Large Language Models (LLMs) like GPT-4. It involves crafting the text or input in a manner that is effectively understood by the AI, leading to the desired output.

4.2.1 Evolutionary Approach to Prompt Development

The development of effective GPT-4 prompts for the PowerGenX project involved a methodical and iterative process. This section outlines the evolutionary stages of prompt engineering, showcasing the progressive refinement from basic summaries to a sophisticated automation blueprint.

1. ***Version 1: Initial Summary Attempt***

   Prompt: "Summarize this document into a PowerPoint outline."
   Result: The output was a rudimentary summary, lacking the structural complexity necessary for effective PowerPoint slides. This initial prompt highlighted the need for more specific guidelines to shape the content into a presentation format.

2. ***Version 2: Enhanced Structure Focus***

   Prompt: "Create a detailed PowerPoint presentation outline from this document, including slide titles and key points."
   Result: The refinement of the prompt led to a more structured outline, aligning closer with PowerPoint slide formats. However, it lacked integration with Python coding for automation, pinpointing the next area of improvement.

3. ***Version 3: Introduction of Automation***

Prompt: "Generate Python code using 'PowerGenX' for a PowerPoint based on this document's structure. Include slide titles and bullet points."
Result: This version produced a more coherent structure accompanied by Python code snippets, meeting basic automation requirements. It demonstrated the potential for integrating text generation with automated slide creation.

4. ***Version 4: Towards a Comprehensive Blueprint***

   Prompt: "Create a concise, engaging PowerPoint presentation outline for a proposal:(input)" The presentation should contain no more than 6 slides, including a title slide. The presentation time is 3 minutes. ", "Provide a title and content for each slide. Generate only Python code (without text) using the following code template as a reference: (input template)"
   Result: The prompt led to the generation of a detailed presentation outline coupled with relevant Python code. It showed significant progress in aligning with the project's automation needs but indicated the necessity for more precise constraints.

5. ***Final Version: Precision and Completeness***

   Prompt (Not Complete): "Create a {text_style}, engaging PowerPoint presentation outline for a proposal. The presentation should have {slide_count} slides, including a title slide. The presentation time is {duration} minutes. Theme: {theme}, Slide Layout: {layout}. Provide a title and content for each slide. Guidelines: [Content]. Reference Code:[Reference Code Provided]"
   Result: Precisely structured presentation outline with Python code fulfilling specific requirements for slide content, format, and style. Demonstrates in-depth understanding of the presentation's needs and optimizes the use of 'Poweregen' for effective PowerPoint creation.

4.2.2 Prompt Engineering Insights

After more than 200 calls to the GPT-4 API to modify the prompt, the final output was stabilized and fulfilled the requirements well. Some important insights: Providing structured guidelines within the prompt, such as slide count, presentation duration, theme, and layout, played an important role in shaping the content into an actionable blueprint for PowerPoint generation. The process emphasized the need for iterative refinement, providing feedback from each prompt's output to fine-tune the subsequent version.

4.3 Image Search and Integration

Concurrently, the system evaluates the necessity for visual content. When image augmentation is required, the system uses the Pexels API to source pertinent images based on the ideas generated by GPT-4. These images are then seamlessly integrated into the presentation code, enriching the slides with relevant visual aids.

4.4 PowerPoint Generator Class

The core of the PowerPoint creation lies within the 'PowergenX Class.' This class, utilizing the 'python-pptx' library, is adept at executing various functions necessary for slide creation. It serves as the backbone, providing a structure upon which the GPT-4 module can append or modify content, including titles, bullet points, and images.

4.4.1 Reference Code for GPT-4

An essential part of PowerGenX's data processing is the reference code provided to GPT-4. This code serves as a learning model for the AI, demonstrating the correct usage of functions in the Presentation Generator Class. The reference code covers a range of functionalities, from basic title addition to complex slide comparisons.

4.4.2 Example

```
a = Powergenx() # Initialize a new Powergenx instance
a.add_title("implementations") # Title slide creation
a.add_content_bullet("implementations") # Slide with bullet point
a.add_content_subbullet("implementations") # Slide with sub-bullet point
a.add_content_bullet("implementations") # Slide with bullet point
a.add_content_bullet("implementations") # Slide with bullet point
a.add_comparison_slide("implementations") # Slide with 2 comparison text boxes
a.add_content_bullet("implementations") # Slide with bullet point
a.save("YourFileName") # Save the PowerPoint
```

4.5 Execution and Output

The result of the process is the execution of the PowerPoint code through the 'run code' function. This function meticulously processes each line of code, invoking the 'PowerGenX Class' methods to render a downloadable PowerPoint presentation. The output is then made available to the user via the Gradio frontend, delivering a PowerPoint presentation alongside an optional script, if requested.

5 Baseline Model

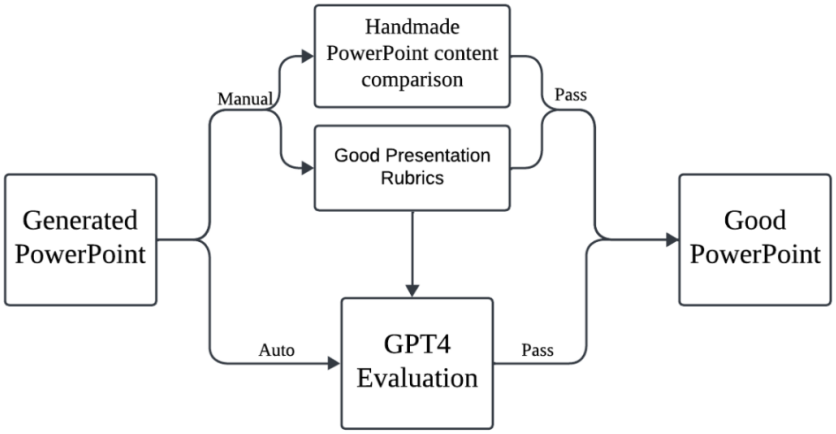

Figure 5: Baseline Model

5.1 Human-Crafted Baseline

The baseline for evaluating the PowerGenX system's output is a PowerPoint created manually, embodying high standards of content organization, design, and clarity—hallmarks of expert presentation crafting.

5.2 Automated GPT-4 Assessment

Concurrently, an automated assessment utilizing GPT-4 offers a quantitative evaluation of the PowerPoint. By inputting the presentation rubrics into GPT-4's interface, the system assesses whether the AI-generated slides adhere to the predefined criteria. This automated process supplements the manual review, providing a data-driven measure of the PowerPoint's quality.

5.3 Outcome Determination

The final quality of the PowerPoint is determined through the convergence of both evaluations. Success is achieved when the AI-generated presentation aligns with the human-crafted baseline

and satisfies the criteria laid out by the automated GPT-4 assessment.

6 Quantitative Results

6.1 Software Functionality and Code Accuracy

The initial quantitative measure of success for PowerGenX is the successful generation of PowerPoint slides without system crashes, with all slides adhering to the specified font sizes and page dimensions. This operational benchmark confirms the reliability of the 'PowerGenX' class within the software architecture. The code accuracy, in this case, is determined to be 100%, as it executes without errors and produces outputs in complete alignment with the input specifications.

6.2 Content Similarity Assessment

For content similarity, a manual comparison is conducted between a PowerPoint crafted based on user-provided requirements and the AI-generated version. The comparison focuses on the content's fidelity to the original document, excluding visual elements.

Figure 6: Comparison 1

As observed, the AI-generated content closely mirrors the human-created version, with minor discrepancies in word choice. The AI version maintains the essence and key points, resulting in a content similarity rating of 100%.

6.3 Visual Content Comparison

In evaluating visual content, the AI's ability to match images from the human-crafted PowerPoint is scrutinized.

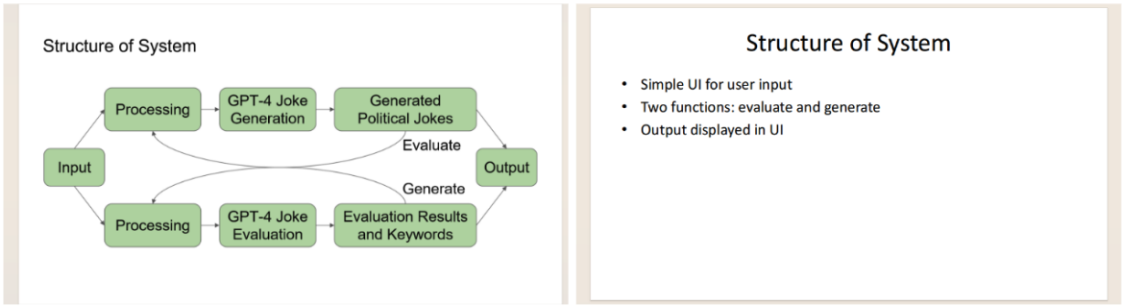

Figure 7: Comparison 2

Although the AI effectively summarizes slide content, it lacks the precision in replicating specific images, capturing only the general concepts. Therefore, the visual content similarity score is reduced to 50%, reflecting the partial alignment with the human-curated graphical representations.

Table 1: PowerPoint Generation Similarity Analysis

PowerPoint Generation	Similarity in Content (Mean Value)
ppt generation 1	100%
ppt generation 2	81.25%
ppt generation 3	58.3%
ppt generation 4	95%

7 Qualitative Results

7.1 Presentation Rubrics and Manual Grading

The qualitative evaluation of PowerGenX's outputs employs a rubric-based approach to ensure that each PowerPoint slide adheres to the established criteria for effective presentation design [2]:

1. ***Conciseness and Clarity:*** Slides should use succinct phrases rather than complete sentences, focusing on essential information only. This aligns with the principle that less is more in presentation design for clear communication.

2. ***Visual Hierarchy and Readability:*** Slides must avoid excessive punctuation and capitalization. Consistent empty spaces and well-organized content enhance readability, while varied font sizes and colors differentiate titles from body text.

3. ***Color Contrast and Aesthetics:*** Effective use of contrasting colors enhances visual appeal and readability. Avoiding patterned backgrounds prevents distraction and maintains focus on text content.

4. ***Labeling and Formatting:*** Titles should be clearly distinguishable with appropriate font sizes (between 35-45 points), and all content should be correctly aligned and formatted within slide boundaries.

5. ***Adherence to the 6 x 6 Rule:*** Bullet points should convey a single idea per line, with a maximum of six words per line and six lines per slide, allowing for minor deviations if necessary.

6. ***Abbreviation Clarity:*** Any abbreviations used should be accompanied by their full forms to avoid ambiguity and maintain clarity.

7.2 Exemplary Presentations

Upon manual review, presentations that meet at least five out of the six established rubrics are deemed successful. Below is an analysis of a presentation that exemplifies this standard. **Input: Progress Report document of PowerGenX.**

- The presentation effectively uses key phrases, avoiding full sentences where not necessary.
- Limited use of punctuation marks, maintaining consistency across the slides.
- A white background with black text ensures high readability.
- Titles are prominently displayed with a font size of 44, adhering to visibility standards.
- Each bullet point succinctly conveys the intended message in under six words.
- The slide content is free of abbreviations, ensuring clarity.

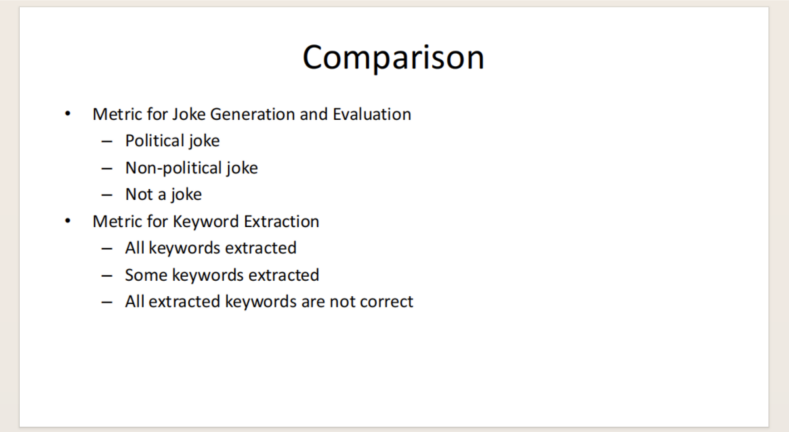

Figure 8: Visual Example of a Good Presentation

7.3 Substandard Presentations

Presentations that fail to adhere to the rubrics are marked down accordingly. The following points detail the shortcomings observed in a lesser-quality presentation:

- Content overflowing the slide borders violates the clarity and space utilization rubric.
- Non-uniform subtitle font sizes fail to maintain a professional and clean appearance.
- With these infringements, the presentation scores a 4/6, categorizing it as inadequate based on our standards.

Figure 9: Screenshots of a 'Bad Presentation'

7.4 Automated Evaluation via GPT-4

The presentations, alongside the rubrics, are also subjected to an automated evaluation by GPT-4. This dual approach helps validate the manual assessments and ensure that the presentations not only look good but also adhere to quality standards algorithmically.

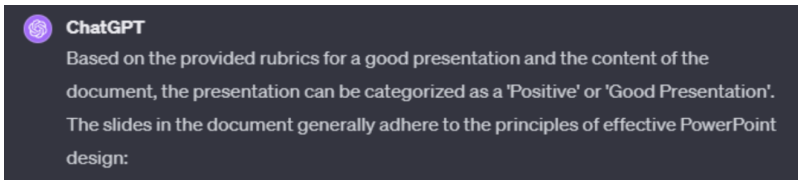

Figure 10: Screenshot from GPT-4 Evaluation

8 Discussion and Learnings

PowerGenX effectively demonstrated the capability of transforming user inputs into structured PowerPoint presentations, showcasing a consistent alignment with the provided documents. A

surprising element was the model's proficiency in adhering to lengthy and complex prompts, accurately following the provided reference code. Despite occasional deviations from the code-only output, the implementation of a specialized function ensured the extraction and execution of the correct code segments, maintaining operational reliability.

Reflecting on the project, we recognized the strengths and limitations of AI in automated content generation. Future projects could benefit from integrating advanced features for creative presentation elements and exploring new input processing methods to further refine automated presentation creation. PowerGenX stands as an example of AI's potential in simplifying content creation, while providing insights into prompt engineering and AI code execution within the realm of current technology.

9 Contribution

Yuquan Gan implemented the whole generator class 'PowergenX' and helped fine-tuning the content prompt of the GPT4 module for code generation. Yuquan Gan also built the Gradio frontend.

Tianze Wang implemented the GPT-4 Interaction Module including prompt development, picture implementation, script creation. Tianze helped to revise the Gradio frontend.

References

[1] Typeset. "Typeset." Accessed: 2023-12-03. (2023), [Online]. Available: `https://www.typeset.com/`.

[2] National Conference of State Legislatures. "Tips for making effective powerpoint presentations." Accessed: 2023-12-03. (2023), [Online]. Available: `https://www.ncsl.org/legislative-staff/lscc/tips-for-making-effective-powerpoint-presentations`.

Permission Table

	Post Video	Post Final Report	Post Source Code
Tianze Wang	No	Yes	No
Yuquan Gan	No	Yes	No