Presentify: Automated Presentation Slide Generation from Research Papers using NLP and Deep Learning (May 2024)

Atul Shreewastav[1], Bidhan Acharya[1], Nischal Paudel[1], and Yugratna Humagain[1]

[1]Department of Electronics and Computer Engineering, IOE, Thapathali Campus

May 13, 2024

Presentify: Automated Presentation Slide Generation from Research Papers using NLP and Deep Learning (May 2024)

Atul Shreewastav (THA077BCT013)[1], Bidhan Acharya (THA077BCT015)[2], Nischal Paudel (THA077BCT028)[3] and Yugratna Humagain(THA077BCT047)[4]

[1,2,3]Department of Electronics and Computer Engineering, IOE, Thapathali Campus, Kathmandu 44600, Nepal

ABSTRACT This study presents a novel approach for automating the generation of presentation slides from academic research papers. We leverage Natural Language Processing (NLP) techniques, particularly a fine-tuned T5 transformer model, to streamline knowledge dissemination within the computer science domain. The model, trained on a custom dataset of research articles, extracts key information and transforms it into well-structured presentation slides. The T5 model's proficiency in text-to-text transfer tasks is enhanced through fine-tuning with domain-specific metrics. This approach enables the automatic conversion of research articles into concise and informative presentation slides. This research has the potential to significantly impact how technical information is summarized and communicated within the research community, fostering greater efficiency and clarity.

INDEX TERMS NLP, T5 model, Gemini model ,Slide generation, Text extraction and Section extraction

I INTRODUCTION

The effective communication of complex information in professional and educational settings often relies on clear and well-structured presentations. However, manually creating these presentations can be time-consuming and resource-intensive, hindering efficient knowledge dissemination. Recent advancements in Natural Language Processing (NLP) offer promising avenues to streamline this process by automating content extraction and organization.

This research addresses the challenges associated with manual slide creation by proposing a novel methodology for the automated generation of presentation slides from academic research papers in the computer science domain. We leverage NLP techniques, specifically a fine-tuned T5 transformer model, to extract key information, summarize content, and generate well-structured presentation slides. This approach aims to empower users, such as researchers, educators, and students, to focus on content creation and knowledge dissemination by automating the time-consuming tasks of information extraction and formatting. Ultimately, this project aspires to contribute to a more efficient and effective presentation creation process, benefiting diverse professional and educational contexts.

II RELATED WORK

Automating the generation of presentation slides from research papers has become an active area of research, with several approaches emerging in recent years. One notable example is "PPSGen, Learning-Based Presentation Slides Generation for Academic Papers" [1] that emphasizes factual accuracy by employing sentence selection algorithms. Another approach, titled "Learning Based Slide Generator" [2] utilizes machine learning techniques to extract relevant information for slide generation . Additionally, research presented in "Technique for Generating Automatic Slides on the basis of Paper Structure Analysis" [3] focuses on analyzing the structure of research papers to generate slides, highlighting the relationship between sections and their content. Furthermore, "DOC2PPT: Automatic Presentation Slides Generation from Scientific Documents" explores automatic slide generation from scientific documents, emphasizing paraphrasing and layout within a hierarchical model [4].

These efforts collectively demonstrate the potential of NLP techniques to streamline the often time-consuming process of slide creation. However, existing approaches face challenges in ensuring the accuracy, relevance, and visual appeal of the generated slides. Our proposed system addresses these challenges by leveraging a fine-tuned

T5 transformer model for content summarization and a dedicated section extraction model (e.g., Gemini model) to preserve the structure and key points of the original research paper.

III METHODOLOGY

A DATASET DESCRIPTION

This research leverages a domain-specific dataset sourced exclusively from arXiv.org, a reputable online repository for computer science research articles. This targeted selection ensures the dataset's relevance to the research focus and provides a well-defined corpus for model training and evaluation.

The dataset consists of 15,655 entries meticulously divided into training (10,958), testing (3,131), and validation sets (1,566) to facilitate robust model performance assessment. Notably, "Computer Vision and Pattern Recognition" emerges as the most prevalent category, with 675 instances.

Each data point comprises two key features: text and summary. The "text" column houses the extracted content from research articles, potentially including sections like introductions, literature reviews, methodologies, results, and conclusions. The corresponding "summary" column contains a concise representation of the extracted text, offering a succinct version of the original content. This dataset serves as the foundation for training and evaluating the employed NLP models. Ultimately, it facilitates the optimization of the system's ability to generate informative presentation slides tailored for academic research papers within the computer science domain.

B PROPOSED METHODOLOGY

This research proposes a novel methodology for automating the generation of presentation slides from academic research papers in the computer science domain. The methodology utilizes Natural Language Processing (NLP) techniques to achieve information extraction, summarization, and presentation structure creation.

1 Text Extraction

The system first extracts the textual content from the PDF file using PyMuPDF, a powerful Python library for working with PDFs. This involves parsing the PDF and extracting the text, including sections, paragraphs, and other relevant information.

2 Section Extraction with Gemini

The extracted text is then passed to Gemini, a tool capable of identifying and extracting distinct sections from the research paper, such as Introduction, Literature Review, Methodology, Results, and Conclusions.

3 Summarization with Fine-Tuned T5 Model

The sections extracted by Gemini are then fed into a fine-tuned T5 model, which transforms the text into concise summaries for each section. The T5 model has been specifically adjusted to optimize the summarization process for research articles.

4 Slide Generation

The summarized content for each section is then passed to the presentation module, which formats the information into a standardized presentation slide format. This ensures the slides have a clear and coherent structure, with each slide covering a specific section of the research paper.

5 Customization Options

The system also provide options for users to customize the appearance and layout of the generated slides, allowing them to tailor the presentation to their specific needs.

C SYSTEM REQUIREMENTS

Following are the basic requirements for the system:

1 Hardware Requirements

- Multi-core CPU (Intel Xeon or AMD Ryzen)
- RAM (16GB or more)
- System Type – 64-bit Operating System

2 Software Requirements

- Operating System – Windows
- RAM Programming Language – Python
- Python Packages - Numpy, Pandas, BeautifulSoup, PyMuPDF

Python

Python is a high level, general purpose, interpreted, dynamic programming language. Python supports multiple programming paradigms, including object-oriented, imperative and functional programming or procedural styles.

NumPy

NumPy is a Python library for numerical computing, offering powerful data structures and tools for working with arrays, matrices, and mathematical functions. Its efficient operations make it essential for scientific computing, data analysis, and machine learning in Python.

Pandas

A Python library specializing in data manipulation and analysis, providing high performance, easy-to-use data structures and tools for working with structured data like tables and time series.

BeautifulSoup

Beautiful Soup is a Python package for parsing HTML and XML documents (including having malformed markup, i.e., non-closed tags, so named after tag soup). It creates a parse tree for parsed pages that can be used to extract data from HTML, which is useful for web scraping. Beautiful Soup

PythonPPTX

PythonPPTX is a Python library used for creating and manipulating Microsoft PowerPoint (.pptx) files programmatically. With this library, you can automate the process of generating presentations by adding slides, inserting text, images, shapes, tables, and charts, and customizing their properties such as formatting, styles, and layouts. It enables tasks like batch processing multiple presentations, dynamically generating slides based on data, and applying design templates for consistent styling. Overall, PythonPPTX streamlines the creation and modification of PowerPoint files, making it a valuable tool for report generation, data visualization, and automated presentation generation.

D SYSTEM BLOCK DIAGRAM

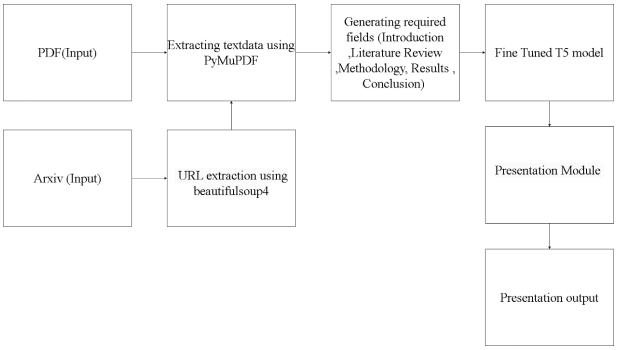

Figure 1: Block Diagram for Presntify

IV RESULTS

A User Interface

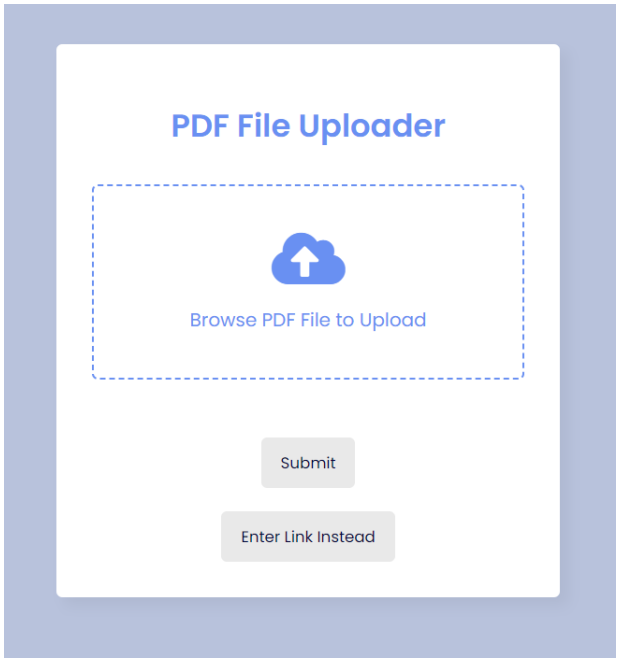

Figure 2: UI for PDF Uploader

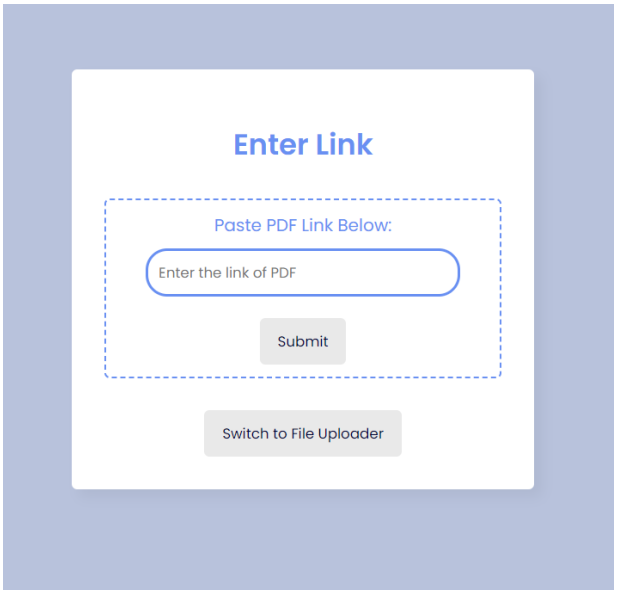

Figure 3: UI for Link Uploader

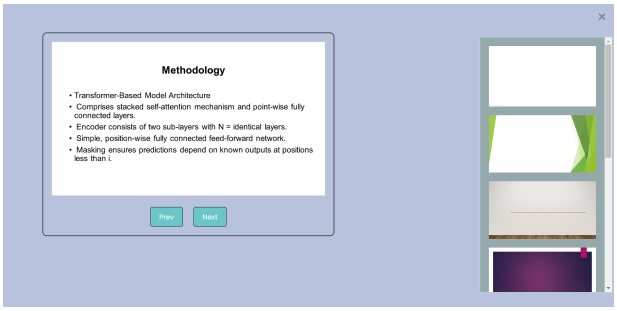

Figure 4: UI for Slide Preview

B Presentation slide result

Upon uploading the research paper ”A Survey on Cross-Domain Sequential Recommendation” the following results were obtained:

Figure 5: Slide 1

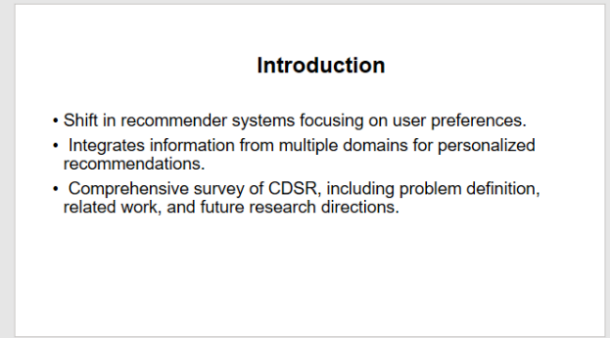

Figure 6: Slide 2

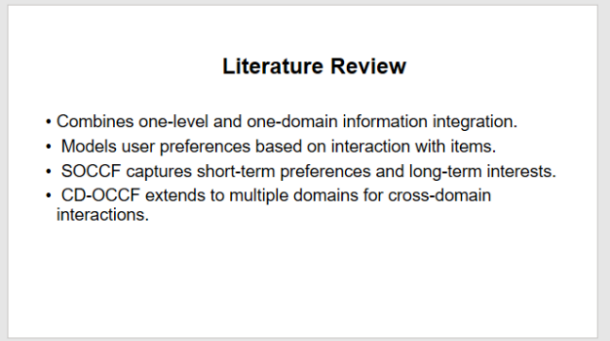

Figure 7: Slide 3

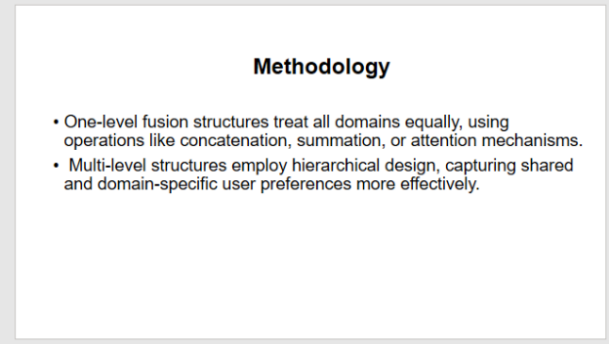

Figure 8: slide4

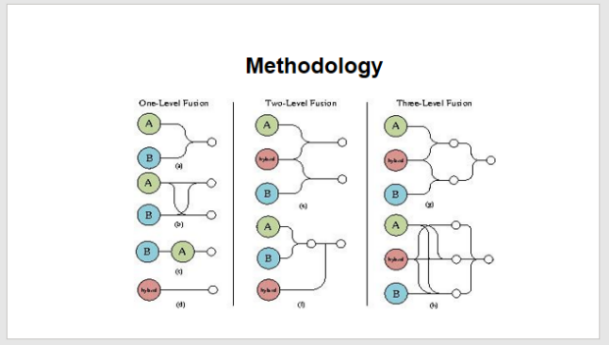

Figure 9: slide5

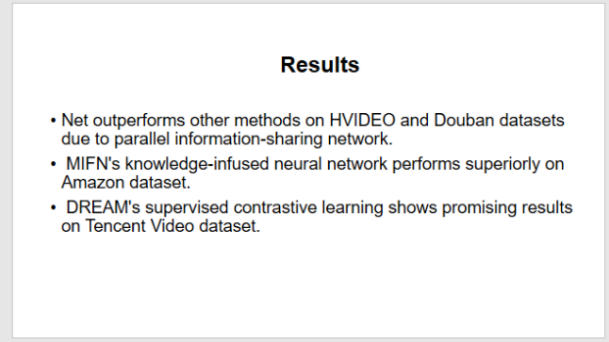

Figure 10: slide6

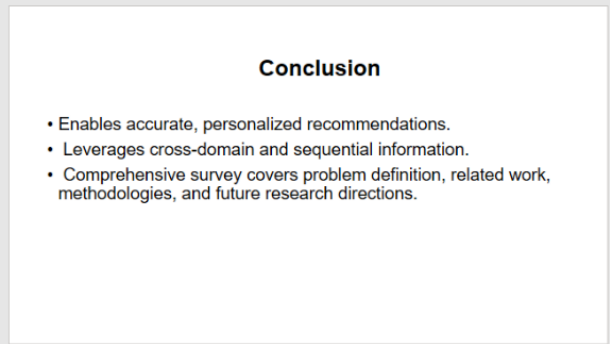

Figure 11: slide7

V CONCLUSION

This research presented a novel tool that utilizes Natural Language Processing (NLP) techniques to automate the generation of presentation slides from academic research papers in the computer science domain. The system leverages a fine-tuned T5 model for content summarization and the Gemini model for section extraction, enabling the extraction of key information from research articles. This information is then transformed into well-structured and visually appealing presentation slides. The system architecture incorporates functionalities for data preparation, model training, text extraction, summarization, and slide generation. The user interface allows for user interaction through features like PDF upload, URL input, slide preview, and theme selection.

The system's performance was evaluated using loss curves and ROUGE scores, demonstrating the effectiveness of the T5 model in summarizing research articles. This evaluation ensures the accuracy and conciseness of the generated summaries. Additionally, the user interface offers a user-friendly experience for researchers, educators, and professionals, empowering them to create informative presentations with greater efficiency.

A ADVANTAGE

- Demonstrates precision in extracting relevant information from documents.
- Efficiently condenses detailed content into impactful slides.
- Versatile in engineering fields.

B DISADVANTAGE

- Projects success depends on Gemini's accuracy.
- Users ability to customize the PowerPoint maybe limited.

- Number of bullets per slide and length of bullet is not configurable by user.
- Resolution and positioning of image is inconsistent.

VI FUTURE SCOPE

This work lays the foundation for a user-customizable summarization system. Future iterations can focus on expanding user control over the front-end experience. This could include implementing a wider variety of themes for visual appeal and allowing users to adjust text size, font, and stylistic elements directly within the website. Additionally, ongoing efforts will prioritize further refinement of the machine learning models used for summarization. This includes continuous parameter optimization, exploration of advanced summarization techniques, and a focus on improving the accuracy, quality, and precision of the generated key insights. By prioritizing accuracy improvements, we aim to deliver a robust system that consistently meets user expectations and effectively addresses their summarization needs.

REFERENCES

[1] Y. Hu and X. Wan, "Ppsgen: Learning-based presentation slides generation for academic papers," *IEEE transactions on knowledge and data engineering*, vol. 27, no. 4, pp. 1085–1097, 2014.

[2] K. Shaj and S. S. John, "Learning based slide generator."

[3] E. Meshram and D. Phalke, "Technique for generating automatic slides on the basis of paper structure analysis," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 5, no. 6, 2016.

[4] T.-J. Fu, W. Y. Wang, D. McDuff, and Y. Song, "Doc2ppt: Automatic presentation slides generation from scientific documents," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 1, 2022, pp. 634–642.

Atul Shreewastav is currently pursuing his undergraduate degree in Computer Engineering at IOE, Thapathali Campus. His research interests encompass various areas, including data mining,automation, and AI/ML.(THA077BCT013)

Bidhan Acharya is currently pursuing his undergraduate degree in Computer Engineering at IOE, Thapathali Campus. His research interests encompass various areas, including data mining, network communications, and optimization theory and technology.(THA077BCT015)

Nischal Paudel is currently pursuing his undergraduate degree in Computer Engineering at IOE, Thapathali Campus. His research interests encompass various areas, including data mining, network communications, and optimization theory and technology.(THA077BCT028)

Yugratna Humagain is currently pursuing his undergraduate degree in Computer Engineering at IOE, Thapathali Campus. His research interests encompass various areas, including data mining, Artificial Intelligence (AI) and Machine Learning (ML) .(THA077BCT047)