Technical Disclosure Commons

Defensive Publications Series

December 2022

User Customizable Content Summarization Using Deep Learning

Joseph Johnson Jr.

Shiblee Hasan

Emmanouil Koukoumidis

David Lee

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation
Johnson Jr., Joseph; Hasan, Shiblee; Koukoumidis, Emmanouil; and Lee, David, "User Customizable Content Summarization Using Deep Learning", Technical Disclosure Commons, (December 13, 2022)
https://www.tdcommons.org/dpubs_series/5584

This work is licensed under a Creative Commons Attribution 4.0 License.
This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

User Customizable Content Summarization Using Deep Learning

ABSTRACT

Users expend substantial time and mental effort in trying to distill the most important facts and/or insight from lengthy text content. While deep learning models can summarize long documents, such models typically operate on the backend. Model output is not customizable by end users via simple user interfaces. This disclosure describes mechanisms that provide users with the ability to obtain customized automatic summaries of documents. Users can specify the desired properties of the summaries based on their preferences and constraints such as available time, desired length or style for the summary, output language, complexity, etc. via convenient UI elements. Users can also choose to obtain automatically generated images or other media that summarize the document content. Customization of output can make the automatically generated summary more relevant and useful and save the user time and effort in reading lengthy text.

KEYWORDS

- Document summarization
- Text summarization
- Content summarization
- Content filtering
- Content simplification
- Deep learning
- Text generation
- Style transfer

BACKGROUND

Users expend substantial time and mental effort in trying to distill the most important facts and/or insight from long documents, such as books, journal articles, white papers, etc. As a result, users may not always come away with the correct takeaways when skimming documents by scrolling through them.

In current applications, users can interact with digital documents via user interface (UI) elements such as sliders and buttons. Users can employ such controls to perform various operations such as changing the text size, zooming in and out, making a page fit on the screen, scrolling through the content page-by-page, etc. However, current applications for reading documents do not provide summaries of the content.

Deep learning techniques can be applied to obtain a summary of a specified length from a longer document of any length. Such an operation can be realized using a trained machine learning model with any suitable architecture, such as a neural network or Probabilities Graphical Model (PGM). Such model architectures are typically based on popular transformer architectures, such as Bidirectional Encoder Representations from Transformers (BERT), Generative Pre-trained Transformer 3 (GPT-3), Pathways Language Model (PaLM), etc. Known word tokens on which the model is trained serve as the input to the model, thus sometimes necessitating the use of domain-specific models.

Increasingly, however, models with large, open-domain transformer architecture are trained with hundreds of billions, or even trillions, of parameters. The learning process for the models encodes each piece of text into a number of embeddings, focusing on the important part of the text specified by an attention mechanism. Less important parts of the text receive much smaller weights in the output encoding representation or are omitted altogether. Based on the

embeddings, the model can output word tokens that can serve as a concise summary of the input document.

Some versions of such models can be implemented to allow user control over the size of the output summary by specifying or limiting the number of words. For example, a user can specify that the generated summary of a given book be limited to a maximum of 2,000 words. Users can specify the desired output parameters via any suitable mechanism that can serve as an input to the model, including natural language phrases (e.g., "Summarize this book in no more than 2,000 words."). For example, such functionality can be used to automate the generation of papers on a given topic.

Similar models are employed to generate images or videos based on user input. For instance, a user can ask for an image of "a Corgi in a doghouse made of sushi" or generate a video from a text description of the desired video content. Conversely, other models can generate a text summary of given video or audio content. However, the current state of the art text summarization capabilities of the various models typically operate on the backend and are not directly available to end users within the graphical user interfaces (GUI) of applications or device operating systems.

DESCRIPTION

This disclosure describes mechanisms that provide users with the ability to obtain automatic summaries of documents via suitable deep learning techniques. For instance, such summaries can condense a multi-page document into a shorter form, such as a few paragraphs, a set of bullet points, a single sentence, etc. Per user preference, the extent of summarization can be based on one or more factors such as the amount of time available to read the summary, maximum permissible semantic loss, etc. For example, if a user has 15 minutes to read a 50-page

chapter of a book, the length and format of the automatically generated summary for the chapter can be adjusted such that it can be read within 15 minutes. Similarly, a student who is reading a textbook or other study material an hour before an upcoming exam can be provided with a summary that can be read and absorbed in less than an hour.

In some cases, users are not constrained by time, but instead wish to obtain help in comprehension of the original content. Such cases can arise, for instance, when the document contains archaic linguistic construction, complex concepts, jargon, or when the document is in a language that is not the user's native language. In such cases, the automatically generated text for the original document can transform the original content to a simplified form without attempting to reduce the length of the content. The model to perform such transformation can be trained on existing resources that involve manual conversion of original content into a simplified form (e.g., "Simple English" Wikipedia). Such a mechanism can additionally employ age-appropriate content filtering for young readers, if age-related information is available and permitted for such use. For instance, the generated text can cleanse sensitive or explicit original content unsuitable for younger readers.

Further, users can be provided the option to adjust the style and tone of the generated summary. For example, a user can ask for a summary of a book in the form of a book review or according to the perspective of utilitarianism. For example, a user can ask for a "250-word negative book review from a transcendentalist." In such cases, an initial summary generated via the deep learning model can be transformed to the desired style and tone via a style-transfer based model trained on the works covering a variety of genres and authors, including those who are no longer alive.

Users can control the properties of the summarization via convenient UI elements, such as sliders, buttons, etc. For instance, users can control the amount of display text with a slider that progressively reduces/increases the amount of text as the user slides left/right, with the leftmost position of the slider providing the shortest possible summary and the rightmost position displaying the original document in full.

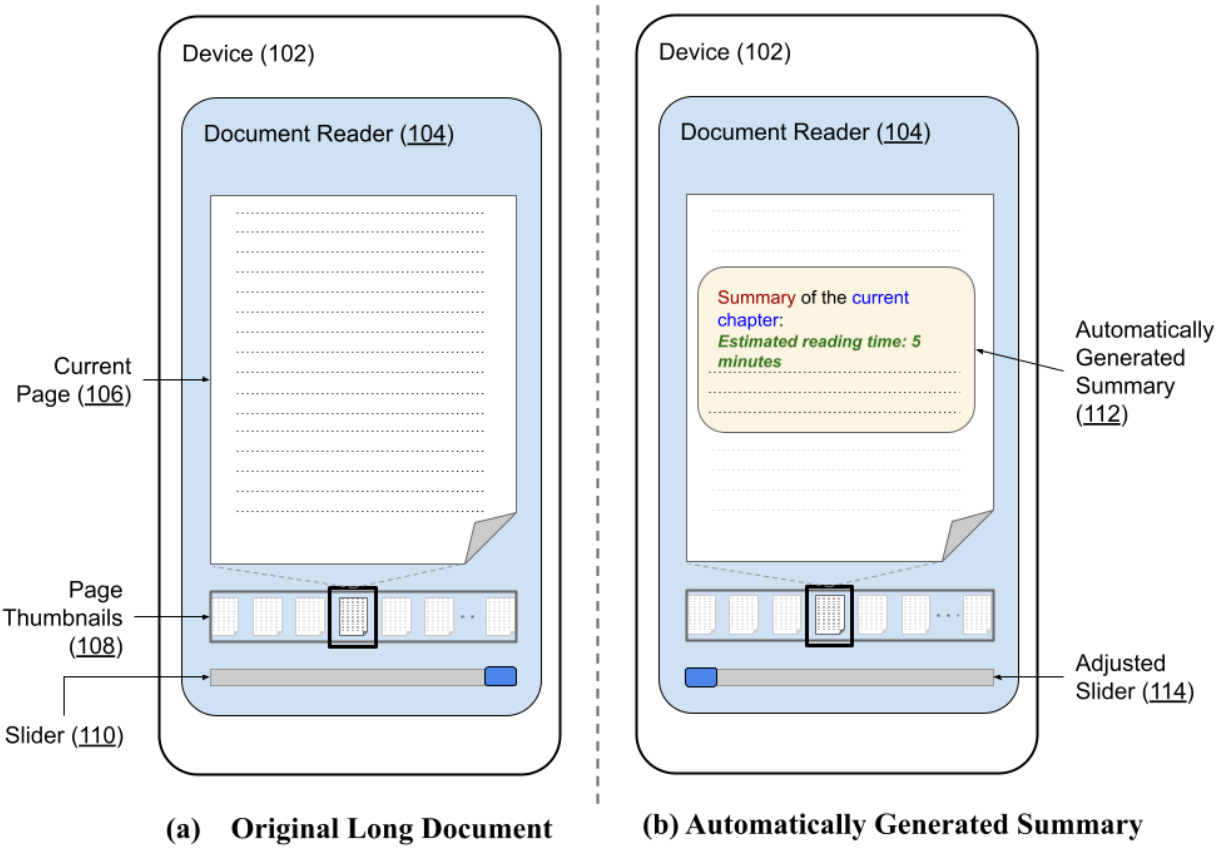

Fig. 1: Controlling automatically generated summaries via a slider

Fig. 1 shows an example operational implementation of the techniques described in this disclosure. A user is reading a long book in a document reader application (104) on a device (102). The application UI includes a slider (110) for the user to specify and adjust the properties of the automatic summarization. In the rightmost position of the slider (Fig. 1a), the user can view the original document as usual with the current page (106) and a set of page thumbnails

(108) that indicate the position of the displayed page within the document. The user can obtain an automatically generated summary (112) by sliding the slider to the left. Moving the slider all the way to the left (Fig. 1b) generates a short summary of the current chapter of the book that can be read within a short time, e.g., 5 minutes.

The user can dynamically adjust the length of the summary by moving the slider between the two ends. The properties of the summary shown at the ends and/or the various intermediate slider positions can be determined by the developers and/or specified by the users and/or inferred dynamically at runtime.

Users can also choose to obtain automatically generated images that summarize the document content. For instance, an image of sailors on a ship at sea can be generated for a book chapter describing a journey on the seas that describes the appearance and temperament of the characters, the weather, etc. Such images can be shown alongside the corresponding text summary.

The techniques described herein can be adapted to operate in the reverse to extend an original document by making it longer with the addition of automatically generated relevant text. For instance, a suitable deep learning model can be applied to generate text that fills in the gaps or clarifies the original text. For example, with author permission, such generated text content can be applied to a novel to refine or add scenes or even characters while maintaining the style, tone, and consistency of the original author. Such an operation can be implemented via a slider that can be moved beyond the position where the original content is shown in full.

The summarization obtained via deep learning models can be augmented in a crowdsourced fashion by aggregating the important parts of the document based on content highlighted by various readers that permit use of their highlights for such purpose. The corpus of

highlights can be used as a training set for a deep learning model to generate automatic highlights in the same manner as text summary. The highlighted text can be connected to a slider that can be moved for progressive highlighting of relatively important content and/or fading of relatively unimportant text. Alternatively, or in addition, the size of important words or phrases can be increased and/or the size of unimportant words or phrases can be decreased so that important text stands out prominently within the document. For instance, single words on a line being larger than every other word on the line can signify the relative importance of the corresponding concept and make the page easier to skim.

The techniques described herein can be implemented for summarizing text documents within any application, device, or platform, such as word processors, document readers, web browsers, e-book readers, etc. In addition, the techniques can be applied for automatic generation of summaries that can be used alongside corresponding content in search engine results, online shopping catalogs, etc. For example, summaries generated via the described techniques can be employed to provide an abstract for an academic paper in the search results for scholarly papers, a review or introductory summary for a book in an online bookstore, etc. With permission, automatically summarized content can be obtained from one or more suitably trained machine learning models and/or other relevant techniques such as crowdsourcing.

While the foregoing description refers to generation of customizable summaries of text documents, the techniques can also be utilized to summarize audio. For example, consider an audiobook that has a duration of two hours. Per the techniques, a summary can be generated of the text of the book and audio can be generated from the summary. The length of the generated summary and audio can be controlled via the slide. Alternatively, a deep neural network can directly summarize and directly generate the audio without using the text.

Implementation of the techniques enables users to control the automatic generation of a summary via backend deep learning models and obtain summaries that are suitable for their preferences and constraints. Users can specify the desired parameters for the summary via any suitable mechanism such as interacting with UI elements, entering text commands, providing voice commands to a virtual assistant, etc. Seamless specification and adjustment of the parameters of the automatically generated summary can make it relevant and useful. Such summaries can save time and effort in reading lengthy text, thus making the user experience (UX) of reading long, complex text content more effective and efficient.

Further to the descriptions above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information (e.g., information about a user's documents, highlights within a document, profession, a user's preferences, or a user's current location), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user, or a user's geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes mechanisms that provide users with the ability to obtain customized automatic summaries of documents. Users can specify the desired properties of the summaries based on their preferences and constraints such as available time, desired length or style for the summary, output language, complexity, etc. via convenient UI elements. Users can also choose to obtain automatically generated images or other media that summarize the document content. Customization of output can make the automatically generated summary more relevant and useful and save the user time and effort in reading lengthy text.

REFERENCES

1. “TLDR this - article summarizer & online text summarizing tool” available online at https://tldrthis.com accessed Nov. 27, 2022.