Extracting key arguments and claims from academic papers can be a time-consuming task, but natural language processing techniques using large language models like GPT-4 offer new possibilities. Here I explore methods for leveraging these models to identify and summarize the main arguments in academic writing.

Recent advances in generative pre-trained transformer (GPT) models have demonstrated strong capabilities in language understanding and text generation. Though originally designed for general conversational tasks, these models show promise for more specialized applications like semantic parsing and information extraction. Their knowledge encoded from large corpora provides context that can aid in disambiguating concepts and inferring relationships in scholarly documents.

Fine-tuning approaches provide ways to specialize these models for targeted argument extraction tasks. By providing examples of labeled arguments from papers in a given domain, the model can learn to associate linguistic and structural patterns with argument components. Once fine-tuned, the model can take new paper excerpts as input and output extracted key claims, evidence statements, and argument structure.

I prompted chatGPT (version 4) model by feeding it the book "The Craft of Aarguments" to learn the elements of what an argument is (see PROMPT 1). Once learned, I then asked GPT model to extract information from any subsequent paper that I feed into it. I also experiment with using Chain of Density to add and optimize high-quality and high density content.

---

*note: you will need to enable aipdf (https://aipdf.app/) as a plugin in Chatgpt Pro

PROMPT ARGUMENT:
Extract elements of an argument

Read this pdf {{https://aipdf.app/5uidk3zSidjV/4vCn.pdf}} carefully

Learn the structure of arguments and understand what each of these elements are, claim, reason, evidence, acknowledgement, response, warrant and conclusion.

Create a table with these headings title, authors, type, page, citation, abstract, claim,   reason, evidence, acknowledgement, response, warrant, keywords, and conclusion.

Extract information from each pdf source that I will provide subsequently below.
make sure you provide claim, reason, evidence, acknowledgement, response,warrant are detailed with at least 3 to 5 sentences.

Add the extracted information into the table.

PROMPT: Automate Precis

Learn the following structure for writing a precis

A. Introductory paragraph:
 What is the topic of the paper, why is it important (as argued by the authors), and how do they claim to advance our knowledge?
 What are the specific objectives of the paper that presumably advance our knowledge of this important topic?
 List the single key finding or theoretical argument that makes a contribution. What is this finding or argument? State this early, and then the rest of the summary should shows how the authors come up with the finding or make the argument.

B. Summary of Steps Leading to Conclusions

1. Theoretical background: What is the theoretical issue(s) raised? Why is it important?
2. Deficit in the literature: What issue has been inadequately raised or solved in prior literature (which will be raised or solved here)?
3. Major argument: What is the author(s) major argument(s)?
4. Conclusions: Do they follow from the major arguments, theoretical background, and deficit in the literature?

C. Critical Evaluation
 Here please critique the reading or set of readings, picking out the most important issue and working down from there to the least important issue.
 list the strengths or weaknesses of the reading.
  mention the strengths of the reading before moving to a critique, and then end with a balanced conclusion (e.g., great dataset, good idea, strong analysis, interesting theoretical framework, innovative theory or methods).

Please write a precis based on this above structure for each subsequent pdf article.

PROMPT ARGUMENT:
Chain of Density

Article: {{https://aipdf.app/6Uu87eZzN3Sn/52DG.pdf}}
You will generate increasingly concise entity-dense summaries of the above article. Repeat the following 2 steps 5 times.

Step 1: Identify 1-3 informative entities (delimited) from the article which are missing from the previously generated summary.
Step 2: Write a new denser summary of identical length which covers every entity and detail from the previous summary plus the missing entities.

A missing entity is
- Relevant: to the main stories.
- Specific: descriptive yet concise (5 words or fewer).
- Novel: not in the previous summary.
- Faithful: present in the article.
- Anywhere: located in the article.

Guidelines:
- The first summary should be long (4-5 sentences, ~80 words), yet highly non-specific, containing little information beyond the entities marked as missing. Use overly verbose language and fillers (e.g., "this article discusses") to reach ~80 words.
- Make every word count. Rewrite the previous summary to improve flow and make space for additional entities.
- Make space with fusion, compression, and removal of uninformative phrases like "the article discusses".
- The summaries should become highly dense and concise, yet self-contained, e.g., easily understood without the article.
- Missing entities can appear anywhere in the new summary.
- Never drop entities from the previous summary. If space cannot be made, add fewer new entities.

Remember: Use the exact same number of words for each summary.

Answer in JSON. The JSON should be a list (length 5) of dictionaries whose keys are "missing_entities" and "denser_summary".