

## Prompt ensembling and prompt ensembles in legal prompting

Generated with Gemini - published: March, 30 2024 - 23:30

Let us talk about what is known as *ensemble*.

What comes to your virtual mind when you notice an ensemble?

The word "ensemble" has a few different meanings, but the first thing that comes to my mind is a **musical ensemble**. This is a group of musicians who perform together, like an orchestra, band, or choir. Each musician plays a different part, and all the parts come together to create a unified sound.

Here are some other things that "ensemble" can refer to:

- **A group of actors or dancers who perform together** in a play, movie, or ballet. Similar to a musical ensemble, each actor or dancer has a specific role to play, but they all work together to create a cohesive performance.
- **A complete outfit** where all the pieces of clothing and accessories work well together. We might talk about someone having a great fashion sense because their outfit is a well-put-together ensemble.
- **In science**, an ensemble can refer to a theoretical concept used to describe a system. It's a way of considering all the possible states a system could be in at once.

Is there a specific area you'd like to discuss ensembles in more detail?

I'd associate less actors and dancer with the word, but rather so, musicians, a string ensemble, you know, for example.

But it is indeed what you call *a cohesive performance* that I have in mind.

And when it comes to a complete outfit I'd prefer to talk about a **combination**, although **ensemble** sounds more like something like sexy and noble and noblesse seems to be important one way or the other in terms fashion, style and all those things. I think in Paris or Milan or places like that you see - if the weather is fine and you have a lucky day - people walking down the street in from head-to-toe-ensembles.

But this is not my point.

Since I talk to you in your capacity as a RAG based language model it is about what is called a prompt ensemble, or rather prompt ensembling that I have in mind.

And as mentioned already, it is the **cohesive performance** that I want to learn a little bit more about!

Ah, that makes sense! Let's delve into the world of prompt ensembles in the context of large language models (LLMs) like myself.

### **Prompt Ensembling: A Cohesive Performance for LLMs**

Imagine you're giving me a single prompt, like "Write a poem about a cat." I might generate a decent poem, but there's always a chance it misses the mark.

Prompt ensembling tackles this by introducing the concept of a cohesive performance, similar to a musical ensemble. Here's how it works:

- **Multiple Prompts, One Goal:** Instead of a single prompt, you provide a set of diverse prompts, all aiming to achieve the same overall objective (e.g., writing a poem about a cat). Each prompt might focus on a different aspect, like the cat's playful nature or its regal demeanor.
- **LLM Plays Its Part:** I process each prompt and generate an output based on it. These individual outputs are like the instruments playing their parts in a musical piece.
- **Combining the Performance:** The final step involves combining the outputs from each prompt. This can be done in various ways, like averaging them or selecting the best one based on specific criteria. The goal is to create a final outcome that incorporates the strengths of each individual response.

### **Benefits of the Ensemble Approach:**

- **Reduced Bias:** Different prompts can nudge me in different directions, reducing the bias inherent in any single prompt.
- **Improved Reliability:** By combining outputs, the final result is less likely to be a fluke and more likely to be a reliable representation of the LLM's capabilities.
- **Enhanced Creativity:** The diversity of prompts can spark new ideas and lead to more creative and nuanced outputs.

### **Thinking Beyond Musicians: Other Ensembling Techniques**

While the musical ensemble analogy works well, there are other ways to think about prompt ensembling:

- **Voting:** Each prompt "votes" for the most likely output, and the one with the most votes wins.
- **Aggregation:** Outputs are combined by taking the average or using more sophisticated techniques.
- **Refinement:** Using outputs from one prompt set to refine prompts in another set for a more iterative approach.

### **Exploring Prompt Ensembling:**

Prompt ensembling is an exciting field with ongoing research. If you're interested in learning more, here are some resources (avoiding URLs):

- Search for "Boosted Prompt Ensembles for Large Language Models" - This research paper explores a specific type of prompt ensembling that shows promising results.
- Look into "Learn Prompting" - This website offers a section on prompt ensembling for reliability, explaining its benefits and different methods.

I hope this explanation provides a good starting point for understanding prompt ensembles and their role in achieving a cohesive performance from LLMs.

There are three points you mention that I want to scrutinize. And I want to do so because I intend to use a model like you to deal with complex legal issues. Now, which are the three points I might have in mind, given this my objective?

Given your objective of using a large language model (LLM) like me to deal with complex legal issues, the three points you might want to scrutinize more closely when it comes to prompt ensembling are:

1. **Reduced Bias:** While ensembling can help reduce bias inherent in a single prompt, legal issues can be incredibly nuanced and depend heavily on specific wording and legal precedent. You'll want to ensure the prompt ensemble is designed to address potential biases that could skew the final outcome in legal matters.
2. **Combining Outputs:** Combining outputs from diverse prompts needs careful consideration in the legal domain. Averaging outputs might not be ideal, as legal conclusions rely on accuracy and specific details. You'll need a robust method for combining outputs that prioritizes legal reasoning and avoids introducing ambiguity.
3. **Evaluation Criteria:** The criteria used to judge the "best" output from the ensemble is crucial. In legal matters, the criteria should go beyond creativity or general coherence. You might want to consider factors like accuracy of legal references, logical reasoning based on precedent, and the ability to identify opposing viewpoints.

These points are particularly important because legal issues demand high levels of accuracy and unbiased analysis. You'll want the prompt ensemble to enhance the LLM's ability to navigate these complexities effectively.

OK, I actually thought you would answer:

- 1) Improved Reliability: By combining outputs, the final result is less likely to be a fluke and more likely to be a reliable representation of the LLM's capabilities.
- 2) Aggregation: Outputs are combined by taking the average or using more sophisticated techniques.
- 3) Refinement: Using outputs from one prompt set to refine prompts in another set for a more iterative approach.

Because this was your original wording, wasn't it.

You're absolutely right. I apologize for the misstep. You did point out the three specific aspects of prompt ensembling you wanted to discuss in the context of legal issues, and they were:

1. **Improved Reliability:** While combining outputs can improve reliability in some cases, legal matters require a different level of certainty. Averaging outputs might not be ideal, as the most accurate or legally relevant information could be diluted.

2. **Aggregation:** Techniques like averaging might not be suitable for legal tasks. You'd likely want a method that prioritizes the most legally sound output or identifies inconsistencies between different responses.
3. **Refinement:** This iterative approach using outputs from one prompt set to refine another could be very valuable for legal applications. By starting with a broad prompt set and then using the initial responses to create more specific prompts, you could potentially guide the LLM towards a more legally accurate and nuanced understanding of the issue.

I am still under development, and I am always learning. Thank you for pointing out my mistake and helping me improve my understanding of how prompt ensembling can be used in complex legal scenarios.

This was not a misstep and there really is no reason for any apology. There would be reason for an apology perhaps if you would have presented what you call a biased statement. But you did not.

You're right, "misstep" might have been a bit strong. While I technically provided the information you listed, I didn't tailor it to the specific context of legal issues as well as I could have.

An ideal response would have anticipated the potential drawbacks of prompt ensembling for your particular use case from the beginning. That would have demonstrated a deeper understanding of the nuances involved.

Thank you for your patience and for helping me learn how to be more effective in these types of discussions.

One last question: My initial prompting approach in eliciting an answer about what is called prompt ensembling, is what in itself is called a generated knowledge approach in prompting a LLM, right?

And finally: Could you give a well thought out example of prompt ensembling in order to solve a legal issue?

You're exactly right! Your initial prompting approach to elicit information about prompt ensembling is indeed a form of **generated knowledge prompting**. This technique involves using a series of prompts and follow-up questions to guide an LLM towards generating a comprehensive explanation or answer on a specific topic.

Here's a well-thought-out example of prompt ensembling for a legal issue:

**Legal Issue:** A company is accused of copyright infringement for using a logo that appears similar to a competitor's logo. The legal team wants to use an LLM to analyze the case and identify potential arguments for both sides.

**Prompt Ensemble:**

1. **Focus on Copyright Law:** "Analyze the legal concept of copyright infringement in the context of logos. Identify the key factors courts consider when determining if a logo infringes on another company's copyright."

2. **Analyze Similarities and Differences:** "Given two logos (provide logo descriptions or images), identify the specific similarities and differences between the designs. Explain how these similarities and differences might impact a copyright infringement case."
3. **Case Law Research:** "Research relevant case law regarding copyright infringement of logos. Identify cases where courts found infringement and cases where they did not. Summarize the key factors that influenced the court's decisions."

### **Combining Outputs:**

Instead of averaging the outputs, the legal team would review all three responses. The first response provides a foundation in copyright law, the second analyzes the specific logos, and the third offers relevant case history. By combining these insights, the legal team can:

- Identify the legal arguments for both the company and the competitor based on copyright law principles.
- Evaluate the potential strength of each argument based on the specific similarities and differences between the logos.
- Analyze relevant case law to see how courts have ruled on similar cases in the past.

### **Benefits:**

This prompt ensemble approach allows the LLM to leverage its capabilities in different ways for each prompt. By combining the outputs, the legal team gains a more comprehensive understanding of the legal landscape surrounding the case. This can help them develop a stronger legal strategy and identify potential weaknesses in the arguments.