TV Show Recommendation App

In this project, I built a tv show recommendation application that uses generative AI to give personalized and accurate recommendations for tv shows. It takes in user info including age, favorite places in the USA, personal background, and favorite shows watched before providing 3 personalized recommendations. In this project, I highlighted few-shot prompting, streamLit, structuring AI pipelines, and AI responsibility.

How to run the project

- 1. Clone the repository
- Make sure you have/add a .env file in the repository that has the openAl api key you will use.
- 3. Create a virtual environment and set up module
 - a. python3 -m venv <venv name>
 - b. source <venv name>/bin/activate
 - c. pip install -r requirements.txt
- 4. Then, you can go into the repository and run streamlit run streamlit app.py

What is in the repository

- requirements.txt libraries needed for application to run
- data/ holds IMDB dataset that was cleaned (eventually was not needed)
- tv_rec_model.ipynb Cleans the imdb dataset and prototype for using few-shot training
- streamlit_app.py main application script to run the service
- utils.py All the helper methods used in the service

How to use the application

1. First enter your age or the age of person you are picking a show for. As AI is integrated into more software each day, we need to make sure we use it responsibly and not immaturely. So if the person is younger than 5 years old, we will not let them use the service. If they are older than that, then the service will aim to only give show recommendations that are age appropriate.

2. Now you have 2 options:

- a. You can click "surprise me" and the system will generate a query that will give you a show to watch. If you like the query, then click "Get Recommendations".
 Otherwise, move on to option b.
- You can answer any combination of the given questions and click "Get Recommendations"

Regardless of your query, the service will then compare 4 different approaches to answer your query:

- Zero-shot prompting with GPT 3.5 turbo This option will use GPT 3.5 turbo to answer your query with no extra context given
- b. Few-shot prompting with GPT 3.5 turbo This option will give GPT 3.5 turbo a few examples of what we are looking for as well as an age restriction and some preferences for the output
- c. Zero-shot prompting with GPT 4 same as 1 but with GPT 4
- d. Few-shot prompting with GPT 4 same as 2 but with GPT 4
- The final option is to request a finer analysis of your request if you want an even more personalized request.
 - After you answer the questions, you can switch from standard mode to structured AI workflow mode
 - b. This mode will go through a structured pipeline that will extract useful information from the query, filter and check recommendations against age restrictions, before finally responding to the user with a clear answer(s) and why each show is a good option given the query (BEST OPTION).

Analysis

Analyzing the model outputs, I observed that few-shot-trained models consistently provide more detailed justifications for their classifications, demonstrating a clearer reasoning process compared to zero-shot models, which often produce inconsistent or less reliable responses.

Also, few-shot models exhibit a stronger ability to tailor their recommendations based on contextual cues, particularly in age-appropriate content moderation. They are more likely to provide recommendations that align with the specified age group or, at the very least, include warnings when suggesting content that may contain mature themes. In contrast, zero-shot models show a tendency to overlook these details, leading to responses that may be less suitable for the intended audience. This highlights the advantage of few-shot learning in enhancing model reliability, interpretability, and ethical safeguards in content recommendation tasks.

Another aspect that I found interesting was that the structured AI workflow pipeline worked significantly better on GPT 3 turbo than zero-shot or few-shot prompting. In fact, the results were detailed and accurate enough that they were better than GPT 4 with zero-shot prompting and comparable to GPT 4 with few-shot prompting. The only downside here was that the recommendations even with the structured pipeline, did not always match in genre (2 out of 11 attempts gave show recs with the wrong genre initially), while GPT 4 with few-shot prompting never got one wrong.

At the same time, the few-shot prompt models sometimes missed details such as recommending a 25 min show when a 40-60 min show was requested, while the structured pipeline almost never made such mistakes.

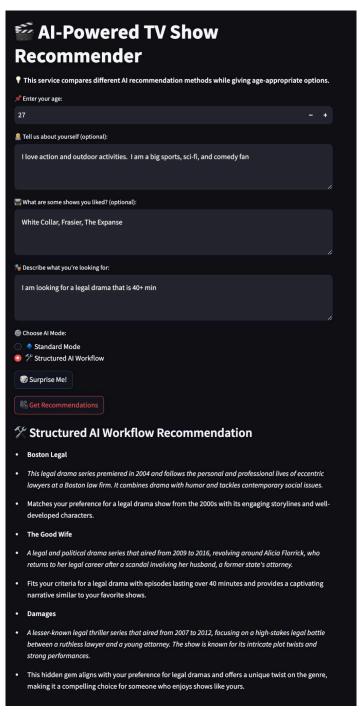
Once I added the additional user inputs and engineered the pipeline prompts to chain better and explain its recommendations, I found the structured pipeline far outperformed any of the models with zero-shot or few-shot prompting. I also optimized the performance vs cost balance by using GPT 3 turbo for the steps in the chain before switching up to GPT 4 for the final output of the pipeline that the user will see.

Workflow progression (including screenshots)

Initially, a user could use only the 4 standard models (GPT 3.5, GPT 4; zero-shot and few-shot) and I gave the user a single prompt to describe what they want:

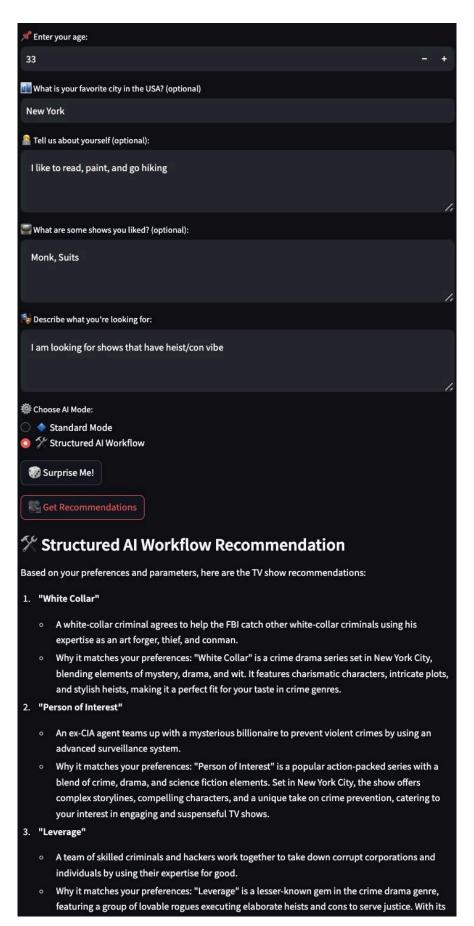


While the few-shot models both gave warnings about the required age to watch the recommended shows, none of them were really geared towards 4 year olds, even though they all gave good recommendations for the query. To improve on this, I built an AI structured pipeline and decided to make it more personalized by adding more user prompts to scrape information from:

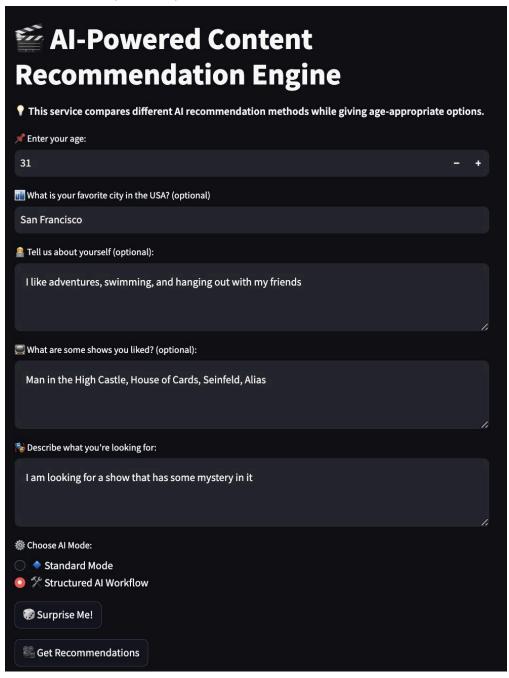


As you can see above, the pipeline combined with the extra information from the user, allows me to not only give better recommendations, but also explain why you might like to watch them (i.e. finding similarities between your favorite shows or activities and the recommended show. Further, to truly personalize the response, I added a nostalgia feature where the first show recommended will be one that premiered around when the user was a child/young adult as they would be more likely to have heard of it but not watched it. The second recommendation is a popular show that the user might like, and the third one is a lesser known show that still fits the user request (if possible), but could be considered a "brand new show" to them. That way, the user gets a variety of shows that fit their interests.

Then, I added a favorite city question to further optimize the recommendations by finding shows that are set in that city or a nearby area (Los Angeles → Southern California; Pittsburgh → greater New England area). While not all of the user information is used in all the recommendations, the questions give enough information for each show recommendation to resonate with the user in some way as you can see below:



Lastly, I added an even more detailed explanation of why each show was chosen for this user so that there is transparency in how AI is being used to personalize your recommendations in each case. Here are a couple examples of the final engine:



Structured AI Workflow Recommendation

Person of Interest

- A high-tech crime thriller centered around a reclusive billionaire software genius and an ex-CIA operative who prevent violent crimes by using an advanced surveillance system.
- Why: If you enjoyed the intricate plotlines and mystery elements in True Detective, Person of Interest offers a captivating blend of crime, drama, and suspense with a unique premise.

The Leftovers

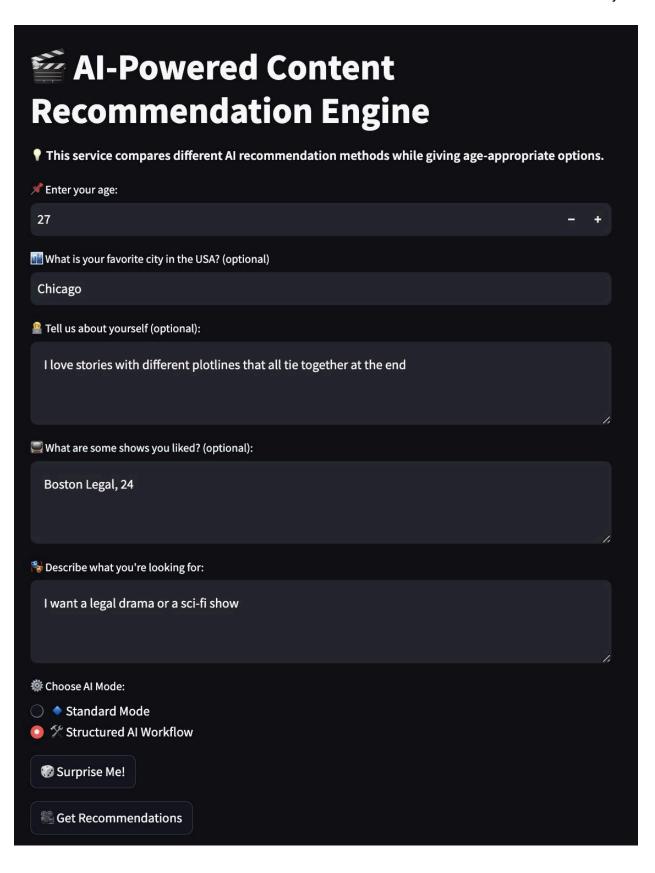
- Set in a small town near New York, this drama explores the aftermath of a global event where 2% of the world's population mysteriously disappears.
- Why: The Leftovers delves into deep character explorations and complex emotional narratives similar to Rectify, offering a thought-provoking and unconventional storytelling experience.

Fringe

- A sci-fi drama following an FBI division investigating unexplained phenomena, including fringe science events that challenge reality.
- Why: If you enjoyed the mix of mystery, adventure, and intricate storytelling in Alias, Fringe offers a similar blend with a touch of science fiction and engaging character dynamics.

Al Explanation: Why These Shows Were Chosen

- Person of Interest: This recommendation is based on the user's love for intricate plotlines and
 elements of mystery as seen in their favorite show, True Detective. Person of Interest, a high-tech
 crime thriller, presents a unique blend of crime, drama, and suspense, which can keep the user
 engrossed. The show's setting, although not in San Francisco, is in New York, presenting an urban
 environment that might be relatable. Moreover, given the user's age, the content is entirely
 appropriate, providing intellectual stimulation and elements of action.
- The Leftovers: As a hidden gem, The Leftovers could be an intriguing recommendation for the user. This drama explores the aftermath of a global event where 2% of the world's population disappears. It offers deep character explorations and complex emotional narratives, similar to Rectify, which is another of the user's favorite shows. The show's mature themes and thought-provoking narratives make it suitable for the user's age. The setting, albeit not in San Francisco, is in a small town near New York, which might resonate with the user due to the similar urban setting.
- Fringe: Given the user's nostalgic decade of the 2000s, Fringe, which aired during that period, would likely invoke nostalgia. The show's mix of mystery, adventure, and intricate storytelling is similar to Alias, another of the user's favorite shows. The touch of science fiction and the engaging character dynamics make



Structured AI Workflow Recommendation

1. The Good Wife (2009-2016)

- This legal drama is set in Chicago and explores the life of a stay-at-home mother who returns to her legal career after her husband's very public sex and political corruption scandal. The show is known for its sophisticated narrative structure, ethical ambiguity, and character development.
- This show matches your preference for a show from the 2000s and also caters to your liking for the city of Chicago.

2. Person of Interest (2011-2016)

- This is a popular sci-fi crime drama that revolves around a former CIA agent who teams up with a
 mysterious billionaire to prevent violent crimes in New York City using an advanced
 supercomputer. The show is well-known for its gripping interconnected storylines and blend of
 legal and sci-fi themes.
- This show is likely to be enjoyed due to its popular nature, complex narratives, and character development, which aligns with your preferences.

3. Damages (2007-2012)

- This lesser-known legal thriller centers on a ruthless lawyer and her recent law school graduate protégé. The show is critically acclaimed for its narrative structure, character portrayal, and the legal and ethical issues it explores.
- Despite not being extremely famous, this hidden gem features intricate plotlines and storytelling similar to Boston Legal and 24, aligning with your preferences.

AI Explanation: Why These Shows Were Chosen

"The Good Wife" is recommended due to its narrative sophistication and character development, two aspects that align with your taste. As a legal drama set in Chicago, it also satisfies your preference for shows set in this city. The story, which revolves around a stay-at-home mother returning to her legal career after her husband's public scandal, is mature and engaging, fitting for your age of 27.

"Person of Interest" is another great recommendation. This sci-fi crime drama, which aired from 2011 to 2016, fits your preference for shows from the 2000s. The show's interconnected storylines and mix of legal and sci-fi themes are likely to captivate you, as they have done for many other viewers. The story, which focuses on a former CIA agent and a billionaire who prevent crimes using a supercomputer, is thrilling yet age-appropriate.

Finally, "Damages" is a hidden gem that matches your liking for intricate plotlines and storytelling. This critically acclaimed legal thriller aired from 2007 to 2012, aligning with your nostalgic decade. The show focuses on a ruthless lawyer and her protégé, exploring legal and ethical issues in depth. Despite not being widely known, its quality and similarity to shows like Boston Legal and 24 make it a perfect fit for your preferences.