

P A R E N T A L

A D V I S O R Y

E X P L I C I T C O N T E N T



Game of Thrones Summarizer

Abhinav Duvvuri, Baisakhi Sarkar, Raagul Nagendran, Swarali Desai



Background



The Problem:

- It's easy to lose track of a show's plotline, especially with busy schedules.
- Googling for summaries often leads to spoilers for future episodes, dampening interest.

The Solution:

- We aim to develop a solution by generating recaps tailored to the user's viewing progress.
- Using subtitle data from all 8 seasons of Game of Thrones, we'll create summaries based on user input.
- Users can also see the top characters and their sentiment change across the selected episodes.



Project Goals:

- Develop a robust model capable of generating accurate summaries based on user input.
- Facilitate catching up with favorite TV shows without encountering spoilers.
- Providing user with top characters for selected range of episodes
- facilitating deeper engagement and understanding of the intricate narrative threads woven throughout each episode

- There is a summarizer exclusively for the Game of Thrones books online, making our tool unique as a summarizer for the Game of Thrones TV show based on user selected episode and seasons ranges.



The background is a highly detailed, aged parchment map from the Game of Thrones series. It features various symbols, including a large circular map in the top left, a star-shaped symbol in the top center, a square map in the top right, and a large black banner with the text 'Data Used' in the center. The map also includes a central castle, a dragon's head, and various other symbols and text.

Data Used

3. Web Scrapped concise summaries and images for each episode of Game of Thrones from IMDb's episode database from

[illegible]



The background is a complex, aged parchment map. At the top left is a large circular diagram with a crosshair. To its right is a sun-like symbol with many pointed rays. Further right is a square diagram with a circular center and the number '33'. At the top right is a large hexagonal diagram with a floral pattern. In the center is a black silhouette of a direwolf's head. Below the direwolf head is a rectangular box containing a paragraph of text in a fictional language. The bottom half of the map features a detailed cityscape with several tall, dark stone towers and spires. The entire map is framed by a decorative border.

Use Cases

User 1: Discontinued the show for a couple of weeks and now wants to get a recap of all the previous seasons

Interaction method: Web application

Wants/Needs:

- Quick summarization of season or episodes to quickly understand or catch up on episodes
- Needs
- Summary of previous watched episodes

Skills: Does not have any technical skills and values a simple interface



User 2: Screenwriter/Content Creators

Interaction method: Web application

Wants/ Needs:

- To research on the show to maybe create a similar show/ prequel or sequel using the episode summary generated.
- More accurate summary

Skills: Does not have any technical skills and values a simple interface



User 3: AI/ML engineer

Interaction method: Interaction with the codebase

Wants / Needs:

- AI/ML engineer who will update the model based on user feedback and other metrics.
- Robust and bug-free model with high availability.
- Access to a dataset of GOT episode transcripts or summaries for training the model.

Skills: Highly technical and knows programming skills.





Component Design



Data Manager

Data Cleaning and Preprocessing:

1. Scrapped concise summaries and images for each episode of Game of Thrones from IMDb's episode database
2. Cleaned data for speaker names and dialogues containing additional punctuations or irrelevant data
3. Merged data for characters to get the top characters based on screen time





User Interface:

- Streamlit Application hosted on Streamlit Community Cloud
- User selects desired Season - Episode range for which they want the summary
- *Sentiment Analysis* of top 3 characters from the episode range selected by user
- *Word Clouds* for the most number of words used by the top 3 characters within the selected episode range
- *Episode Summary*: Summary of the selected range of episodes chosen by the user, along with accompanying images extracted from the episodes





Deployed Model:

1. **Summarizer:** A zero-shot GPT-3.5 model hosted on Azure.
2. **Word Cloud:** Used TF-IDF for picking important words in data.





The background is a highly detailed, aged parchment map from the Game of Thrones series. It features various symbols, including a large circular emblem with a cross, a sun-like symbol with a face, a castle, and the Iron Throne. The map is framed by a decorative border. The title "Lessons Learnt and Future work" is overlaid on a dark, semi-transparent rectangular box at the bottom.

Lessons Learnt and Future work

Lessons learnt

- Having comprehensive test suite to detect bugs
- Pair programming >>>> working alone

Future work

- Convert the GPT model to a few-shot model and ground it to the transcripts of the show.
- Make a generalized tool to get summaries for any TV show the user is interested in
- Recommend similar episodes from other TV shows based on user selection





Valar Morghulis



One runs means need followed up with the most of the others. you could back to the last
Hale. In the way, send the. Who Liden for the first. omgung rids is not. That is, the
here is got sunal siffat comdrens has hast must out for Mithers, hote horetren hons.
Difan comes off! "Spore not ouling, and its and thide is me gromes the Perylinn mra.

Demo