Information extraction from the given video:
1. Video Captioning
2. Transcription

Video Captioning:
1. Breaking the video into smaller clips
2. Using timesformer-gpt2-video-captioning Pre trained model.
   a. While experimenting, we inferred that It is extracting 8 key frames and doing image captioning
   b. In order, effectively capture more frames, we broke the video into smaller clips and used the model.

Reference for video captioning used:
https://www.analyticsvidhya.com/blog/2022/01/youtube-summariser-mini-nlp-project/;

Story Generation using Audio Transcript :
1. We extracted transcripts from the video using YouTubeTranscriptApi
2. Then we used Mixtral-8x7B-v0.1 Pre trained Model, to generate the story based on the plot_tags and Audio Transcript.