Multilingual Video Synopsis Generator

Abstract

The **Multilingual Video Synopsis Generator** is an AI-powered system that transcribes educational videos, summarizes key content, and translates it into multiple languages—making learning faster, more accessible, and inclusive for global audiences.

Process (Pipeline)

The system follows a structured pipeline to process educational videos and produce multilingual synopses:

1. Video Input – Users can either upload a video file or provide a link (e.g., YouTube).

2. Audio Extraction – The system automatically extracts the audio track from the video.

3. Transcription – A speech-to-text model (e.g., OpenAI Whisper or Google Speech-to-Text) converts the spoken words into English text.

4. Summarization – The transcript is passed through a summarization model (T5, BART, Pegasus) that condenses the content into a coherent synopsis.

5. Translation – The synopsis is translated into multiple languages using machine translation models (Helsinki-NLP or Google Translate API).

6. Output – The final multilingual synopsis is presented with options to read, download, or listen via text-to-speech.

Learning Technologies

The Multilingual Video Synopsis Generator leverages state-of-the-art technologies in artificial intelligence and software development:

• Speech-to-Text: OpenAI Whisper, Google Speech-to-Text API.

• Summarization: Hugging Face Transformers (T5, BART, Pegasus).

• Translation: Helsinki-NLP models, Google Translate API.

• Backend: Flask / FastAPI for handling API requests and pipeline orchestration.

• Frontend: HTML, CSS, JavaScript, Tailwind CSS for responsive design.

• Optional: Text-to-Speech for audio playback and PDF/TXT export for summaries.

Conclusion

The **Multilingual Video Synopsis Generator** revolutionizes digital education by making learning **faster, smarter, and globally accessible**. Through transcription, summarization, and translation, it empowers learners to save time, overcome language barriers, and engage with educational content more effectively across diverse contexts.