Proof Of Concept (POC)

Project Title:

Predicting the Programming Language through Video and Audio Content

About:

It talks specifically regarding the verification and prediction of a programming language through the contents of a video, especially through audio as well as visual analysis.

Focus Areas:

Audio transcription: Extract and transcribe audio file content using a speech recognition model from the video and store the text.

Video transcription: Extract the information from the video such as code snippets, information, etc using OCR models

Programming language prediction: prediction of programming languages using the text transcribed from the audio and the code snippets from the video.

Sentiment Analysis: running sentiment analysis on the transcribed text to check whether there is any inappropriate or offending content. Video Frame Analysis is performed using rekognize model from aws to validate the relevance of the video and avoid inappropriate content

Synchronization Check: The video will be rejected if the audio and video contents doesn't match.

Significant Features:

Audio Recording: The Whisper model, developed by OpenAI, transcribes the audio of the video. The transcription is sent for more complex analysis, such as sentiment and programming language prediction.

programming language prediction: The model should predict the programming language by matching commonly associated keywords with the programming language, such as Python, Java, JavaScript, C#.

For example, words like "class," "function," or "Flask" indicate Python, and words like "Spring" or "JRE" may indicate Java.

Sentiment Analysis: For tone and sentiment analysis, the project applied the feature in the nltk library called SentimentIntensityAnalyzer.

Thus, this analysis can detect whether this video contains any offending or inappropriate language.

Proof Of Concept (POC)

Frame Extraction and Code Detection: Video frames extracted will be used to achieve code detection, thereby further refining the correctness of determining whether the video contains relevant material concerning programming. This part of the project shall be a placeholder and implemented by making use of image processing and OCR tools.

Audio-Video Synchronization: This helps harmonize the translated audio and the visual content to prove the pertinence and coherency of the content being presented.

<u>Workflow:</u>

- A video file full of coding lessons, or programmes-related content.
- Audio Processing The audio is extracted from the video and transcribed using Whisper.
- Programming Language Prediction: This is done by analyzing the transcribed text for prediction of the programming language itself using predefined keyword lists.
- Sentiment and Content Analysis: Sentiment analysis is used for discerning whether the video contains any inappropriate language.
- Video Frame Analysis-Optional The video frames are run over to check for the presence of visible code.

<u>Example Cases:</u>

A video on the basics of Python passes through the system, detecting related keywords such as "Flask," "Django," "data science," and so forth.

A Java program is being video-processed, after which it identified the terms "Spring," "object-oriented," and "JDK," wrapping up the topic-which is about a talk of Java.

The transcribed audio does not contain any offending content.