

Documentation for building Speech recognition

Sunday, 5 November 2023 3:42 pm

Tech stack: Streamlit
Huggingface pipeline

IDE: Visual Studio Code

Model used: "distil-Whisper/distil-large-V2"

Script location: GitHub link

Hosting: [huggingface link](#)

Challenges

Sunday, 5 November 2023 4:42 pm

How to implement caching:

Initially implemented using `st.cache` but was deprecated

The solution was found at streamlit docs caching resources using `st.cache_resource` to cache the model, this is to prevent the model from loading on every run, thus reducing run time.

How to include file upload:

Use streamlit file upload however the returned class cannot be processed by the huggingface pipeline

The solution was found in the streamlit forum. The solution was to use `tempfile` library `Namedtempfile` to store the recorded audio as a file temporarily before passing it to the speech recognition model. Makes processing slow

How to catch errors.

The solution was to use `try` and `except`

How do I get users to record audio:

The solution was found using the streamlit component(`streamlit-audiorec`)

The app is still slow:

After implementing streamlit caching and other techniques to reduce run time the app is still slow because it's running on a slow compute unit (16 vCPU)