

FINE-TUNING WHISPER FOR SPEECH TRANSCRIPTION

Trelis Research

OVERVIEW

1. WHISPER DEMO!
2. USES CASES FOR FINE-TUNING A TRANSCRIPTION MODEL!
3. HOW TRANSCRIPTION WORKS!
4. PREPARING TO FINE-TUNE
5. REPO OVERVIEW
6. NOTEBOOK DEMO!
7. PRO TIPS!



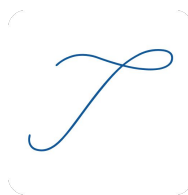
FINE-TUNING USE CASES

1. ADD NEW VOCABULARY (NEW WORDS, PHRASES)
2. IMPROVE PERFORMANCE ON ACCENTS
3. IMPROVE PERFORMANCE ON UNCOMMON LANGUAGES



HOW TRANSCRIPTION WORKS!

1. SOUND IS A VIBRATION - MEASURE THAT VIBRATION!
2. CONVERT THE VIBRATION INTO FREQUENCIES (THAT HUMANS HEAR)
3. INPUT THOSE FREQUENCIES AS INFO TO A LANGUAGE MODEL ("FEATURES")
4. GENERATE TOKENS (SUB-WORDS)



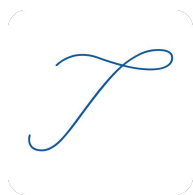
PREPARATION FOR FINE-TUNING

WE NEED :

- A. SOUND (AN MP3 FILE)
- B. TEXT (A TRANSCRIPT, IN VTT FORMAT)

WE NEED THIS FOR :

- TRAINING DATA
- VALIDATION DATA



FINE-TUNING WITH LORA

