

Main Goal

Develop a complete Product:

Subtitle Generator with Speech Recognition

Problem:

1. **Little** AI / Web Develop / Speech Recognition Experience – Both of us

1.1 Detail info (What we know):

AI – Python Sklearn dealing with sentiment data analysis

Web – Flask

Speech Recognition – Zero Experience

Strategy:

1. **Mainly focus** on developing a **complete product** with Baidu API

1.1 Function need to be implemented

Core:

API Interaction – Sending audio, Get result

Subtitle related – Sperate Audio into piece, time recording, output

Additional:

Language Support

Ability of processing video / various types of audio

Realtime output

Make it a web application

Schedule

Workable before late March (Around 3/22)

3/22 – Early May: Try to build our own model (See Below)

After that - Challenging:

Make best effort to build our own model

Potential Solution:

Mozilla Deep Speech

1. Open Source
2. With enough guide on GitHub

Amazon LightSail

1. For training model
2. Maybe Web App

Training:

1. Free dataset provided by TensorFlow

Problem

1. Zero experience – both of us
2. Not sure we can make it or not
3. Seems that Amazon LightSail doesn't provide GPU support (need to upgrade the service)

- a) Hardware problem – Only CPU, should be enough for training model

Linux/Unix		Windows	
\$3.50 USD/mo		\$5 USD/mo	
512 MB Memory 1 Core Processor 20 GB SSD Disk 1 TB Transfer*		1 GB Memory 1 Core Processor 40 GB SSD Disk 2 TB Transfer*	
\$10 USD/mo		\$20 USD/mo	
2 GB Memory 1 Core Processor 60 GB SSD Disk 3 TB Transfer*		4 GB Memory 2 Core Processor 80 GB SSD Disk 4 TB Transfer*	
\$40 USD/mo		\$80 USD/mo	
8 GB Memory 2 Core Processor 160 GB SSD Disk 5 TB Transfer*		16 GB Memory 4 Core Processor 320 GB SSD Disk 6 TB Transfer*	
\$160 USD/mo		\$160 USD/mo	
32 GB Memory 8 Core Processor 640 GB SSD Disk 7 TB Transfer*			

- b) Develop environment – Should be **Linux**

4. Any recommend material / guide?