-Tensorflow를 이용하여 학습.

-IEMOCAP, RAVDESS에서 제공하는 dataset을 이용하기로함.

Iemocap==총 12시간의 분량으로 10명의 배우(5명 여자, 5명 남자)가 짝으로 5session을 진행한다. 남녀 대화로 진행이 되고, 모든 대화는 적절한 감정이 포함된다. 각 문장의 emotional label은 사람의 주관적인 평가에 기초하여 할당되어 있다. Script, Improvise session 두 가지로 구성되어 있는데, script의 경우 배우들에게 대본을 주어 recording, improvise의 경우 상황만 배우들에게 주어 즉흥적인 대화를 recording한 것이다. *wav*형식의 데이터이다.

Ravdess==24명의 다른 배우들로부터 녹음된 1500개의 파일. (여자12, 남자12), 8개의 다른 감정을 포함.

-감정 카테고리를 8개로 한정하여 각각 번호를 지정하여 구분하기로 함.

-My Mood Music을 예로 들면 감정 카테고리 1번을 100번학습 후 2번을 100번학습…으로 하면 학습이 안된다. 1-8번까지 전부 집어넣고 100번을 학습해야 올바른 W가 나옴.

-CNN(합성곱 신경망)을 사용.

-dictionary형태로 집어넣으면 에러 발생, npz형태로 집어넣어야 함.

-librosa를 이용하여 벡터를 추출.

-히든 레이어 3개(300개->200개->100개)를 이용한다. 시그모이드, 하이퍼볼릭탄젠트, 렐루 중 성능을 비교하여 적절한 것을 배치. 히든 레이어 개수도 조정해본다.