

# HLT Homework 1

Mou Zhang

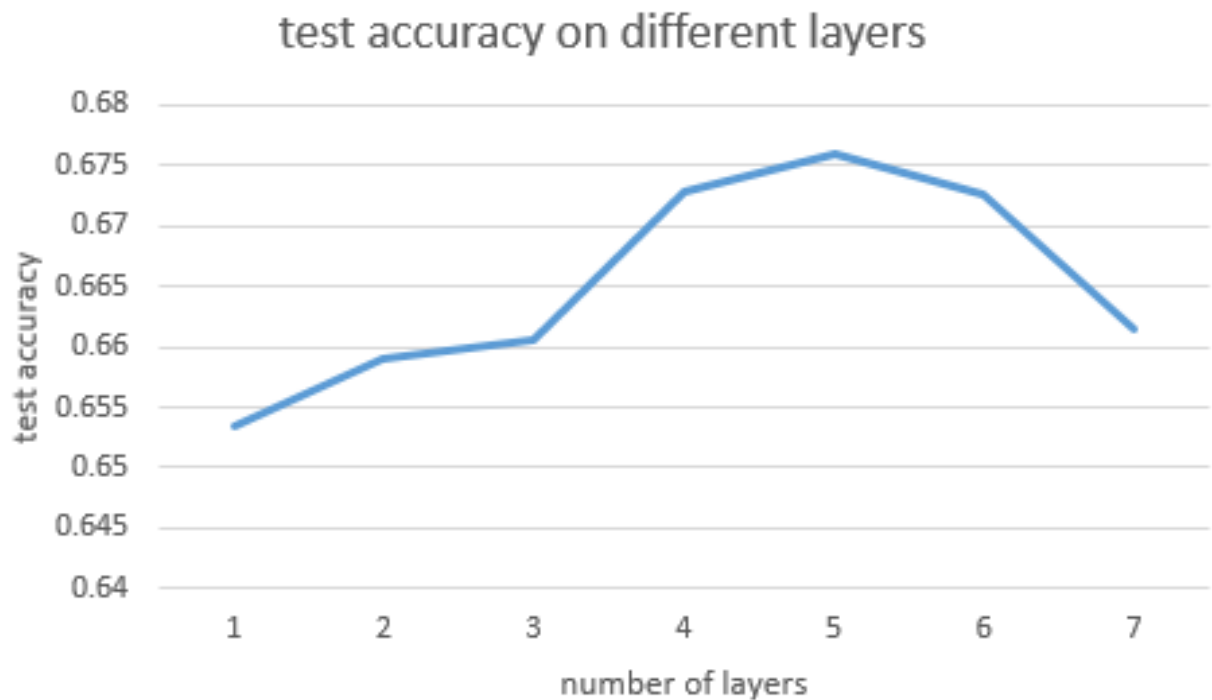
October 2019

## 1 Question 1

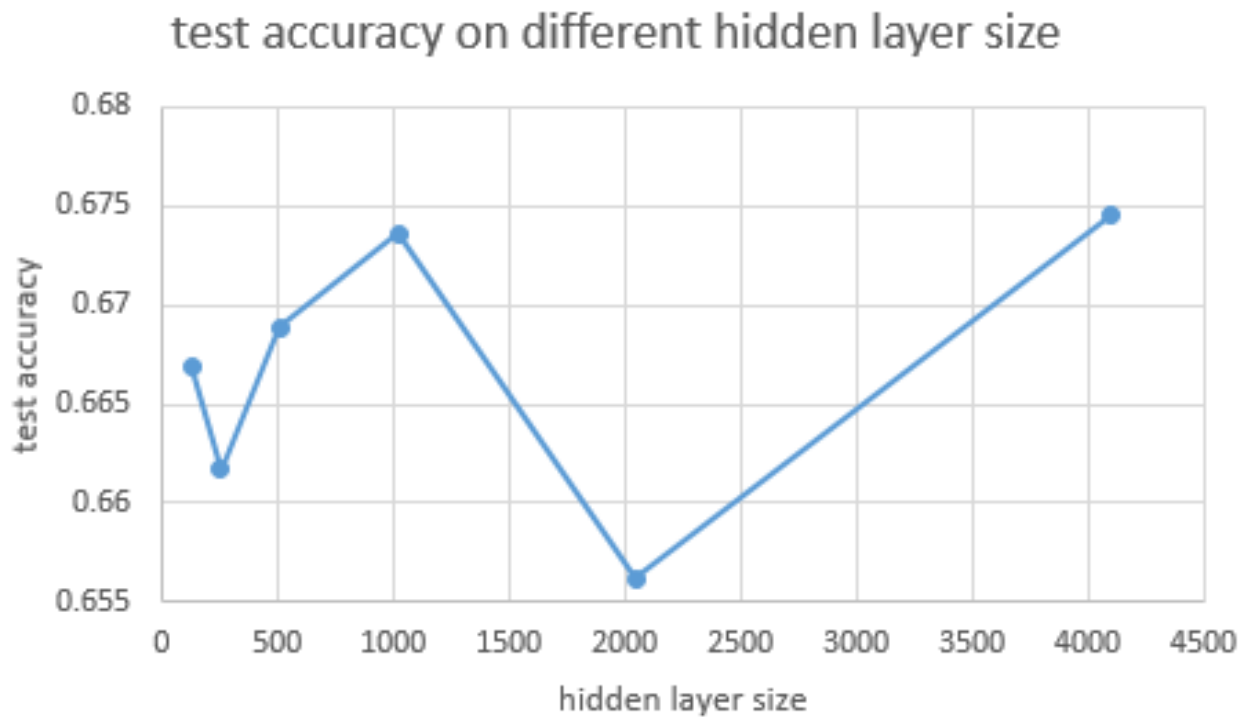
- [Mou Zhang's Public-phoneme-recognition.ipynb](#)

## 2 Question 2

- Here's the the result of test accuracy on different num of layers:



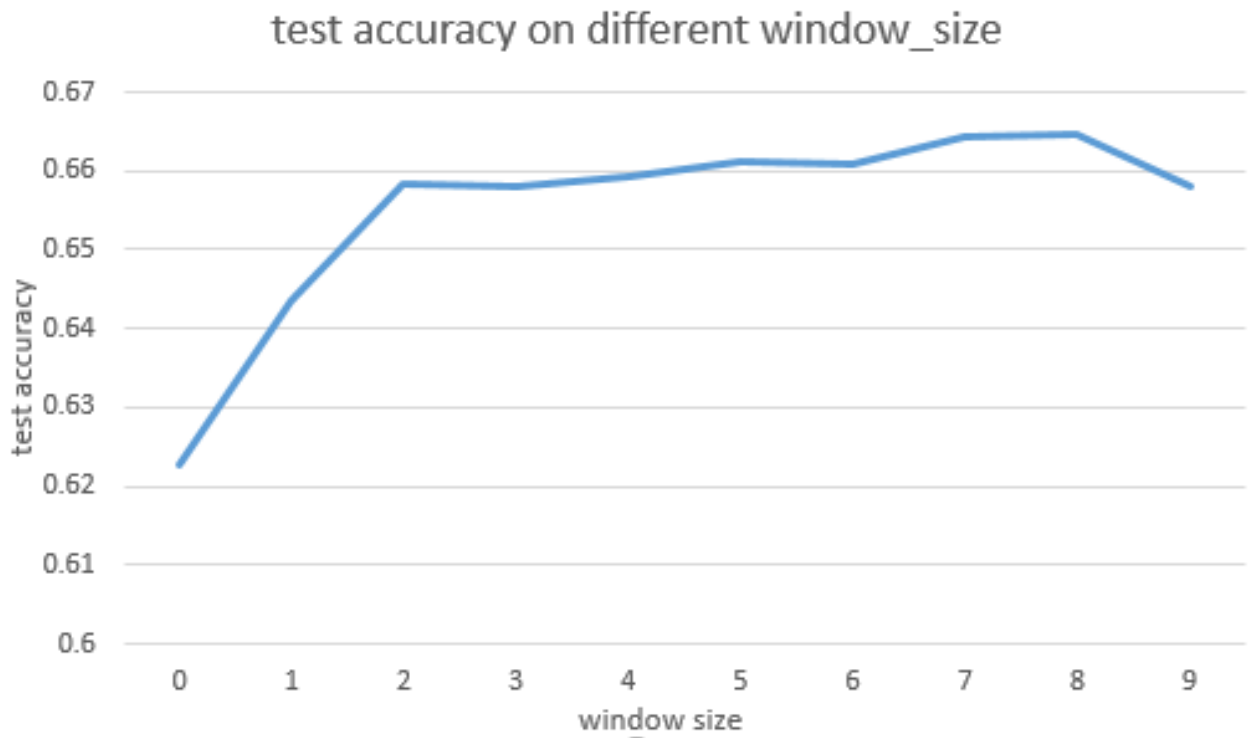
- As we can see from the picture, 5 linear layers works the best.
- Here's the result of test accuracy on different hidden layer sizes:



- From my result, I find out that a hidden size of 4096 works the best, and a hidden size of 1024 works the second best.

### 3 Question 3

- Here's the the result of test accuracy on different window sizes:



- As we can see from the picture, A window size of 8 works the best. Increasing the context really

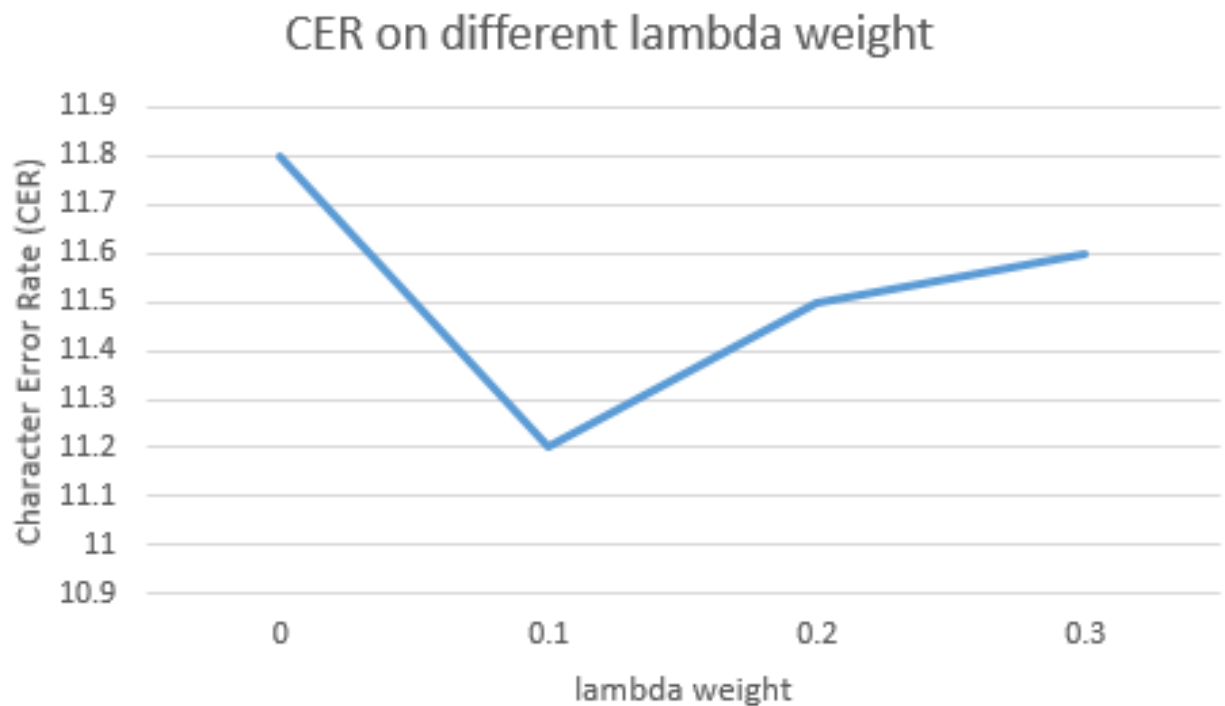
increases the accuracy.

## 4 Question 4

- [Mou Zhang's e2e ASR homework v1](#)

## 5 Question 5

- Here's the the result of test accuracy on different lambda weight:



- As we can see from the picture, 0.1 lambda weight works the best. The language model fusion helps to decrease the CER. But increasing the lambda weight doesn't always help with the CER when lambda increases.