Machine Learning HW2

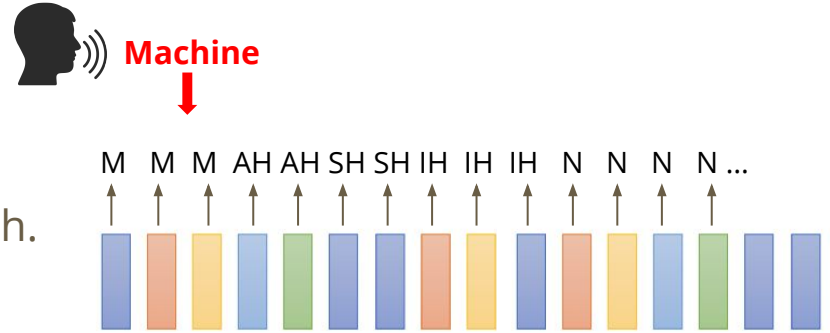
ML TAs
ntu-ml-2021spring-ta@googlegroups.com

Outline

- 2-1 Phoneme Classification (8pt/10pt)
  - Task Introduction
  - Dataset & Data Format
  - Kaggle Submission Format
  - Grading
- 2-2 Hessian Matrix (2pt/10pt)
  - Task Introduction
  - Gradient Norm / Minimal Ratio
  - NTU Cool Submission
  - Grading

2-1 Phoneme Classification

Task Introduction



Task: Multiclass Classification

Framewise phoneme prediction from speech.

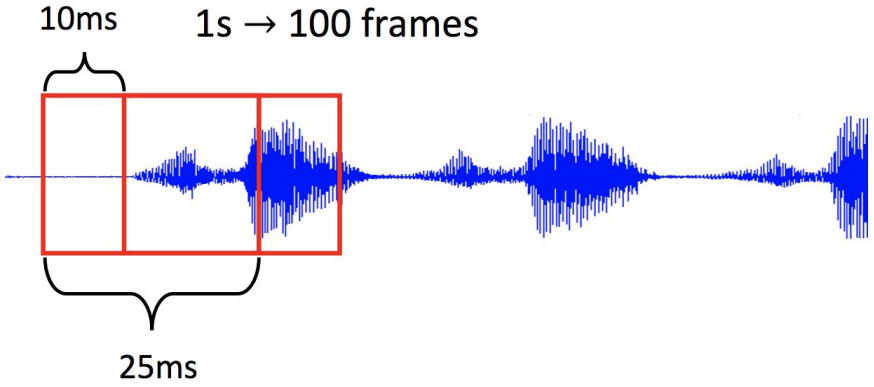
What is a phoneme?

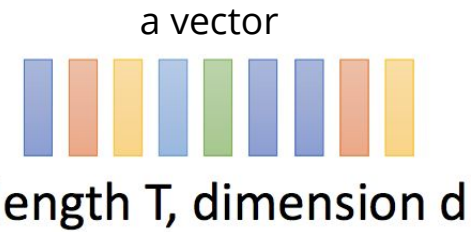
A unit of speech sound in a language that can serve to distinguish one word from the other.

- bat / pat , bad / bed
- Machine Learning → M AH SH IH N L ER N IH NG

Task Introduction

Data Preprocessing



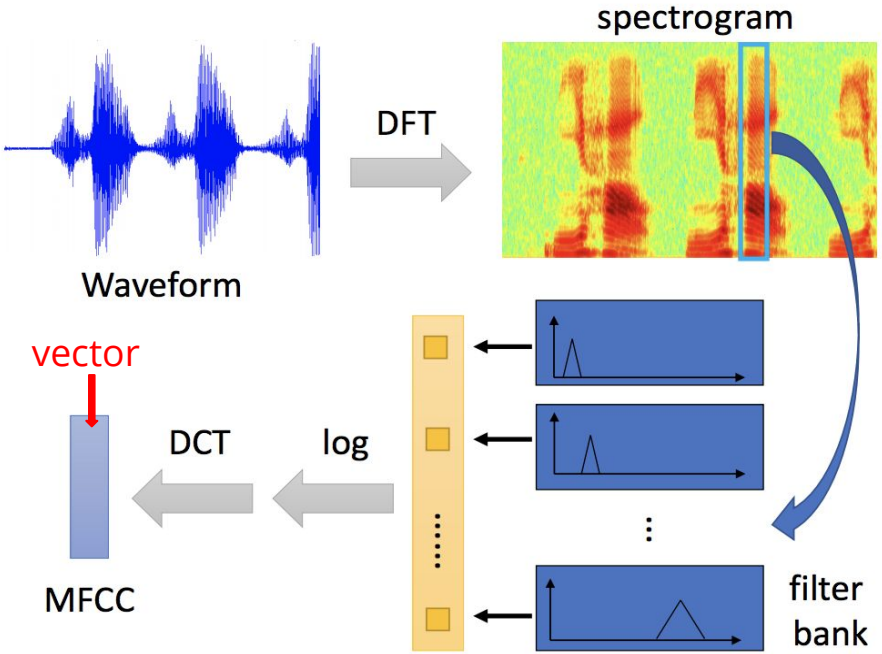


frame

39-dim MFCC
80-dim filter bank
...

Task Introduction

Acoustic Features - MFCCs (Mel Frequency Cepstral Coefficients)



For more details,
please refer to Prof. Lin-Shan Lee's
[Introduction to Digital Speech Processing] Chap.7

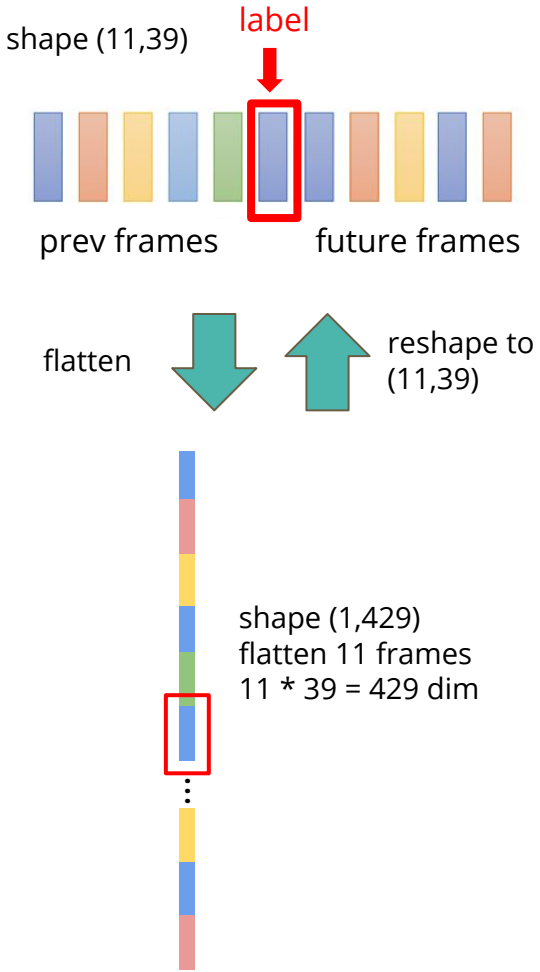
Image ref.
Prof. Hung-Yi Lee
[2020Spring DLHLP] Speech Recognition

More Information About the Data

Since each frame only contains 25 ms of speech, a single frame is unlikely to represent a complete phoneme

- Usually, a phoneme will span several frames
  - Hint: post-processing may help
- Concatenate the neighboring phonemes for training
  - In this HW, we concatenate the past and the future five frames for training (total 11 frames)
  - You may reshape the input (1,429) back to (11,39) to get separated 11 frames
  - Just remember that the label corresponds to the center frame
- Finding testing labels or doing human labeling are strictly prohibited!

Introduction to Digital Speech Processing



Dataset & Data Format

- Dataset: TIMIT Acoustic-Phonetic Continuous Speech Corpus
  - Phonetically balanced for English
- Data Format (The TAs have already preprocessed the data)

  timit_11/
  - **train_11.npy → training data (# of training frames, 11 x feature dim)**
  - **train_label_11.npy → framewise phoneme label (0-38)**
  - **test_11.npy → testing data (# of testing frames, 11 x feature dim)**
  - Acoustic features (39-dim MFCC)
    - Concatenate the past and the future five frames (feature dim = 11 x 39)
    - The phoneme label of each input corresponds to the center frame
- **Using additional data is prohibited**. Your final grade will be multiplied by 0.9!

Class	Phoneme	Example	Class	Phoneme	Example	Class	Phoneme	Example
0	iy	b***ee***t	13	l	***l***ay	26	dx	mu***dd***y
1	ih	b***i***t	14	r	***r***ay	27	g	***g***ay
2	eh	b***e***t	15	y	***y***acht	28	p	***p***ea
3	ae	b***a***t	16	w	***w***ay	29	t	***t***ea
4	ah	b***u***t	17	er	b***ir***d	30	k	***k***ey
5	uw	b***oo***t	18	m	***m***om	31	z	***z***one
6	uh	b***oo***k	19	n	***n***oon	32	v	***v***an
7	aa	b***o***b	20	ng	si***ng***	33	f	***f***in
8	ey	b***ai***t	21	ch	***ch***oke	34	th	***th***in
9	ay	b***i***te	22	jh	***j***oke	35	s	***s***ea
10	oy	b***o***y	23	dh	t***h***en	36	sh	***sh***e
11	aw	b***ou***t	24	b	***b***ee	37	hh	***h***ay
12	ow	b***oa***t	25	d	***d***ay	38	sil	silence/closure sounds

Sample Code

Colab Link:
https://colab.research.google.com/github/ga642381/ML2021-Spring/blob/main/HW02/HW02-1.ipynb

- Simple baseline
  - You should able to pass the simple baseline using the sample code provided.
- Strong baseline
  - Model architecture (layers? dimension? activation function?)
  - Training (batch size? optimizer? learning rate? epoch?)
  - Tips (batch norm? dropout? regularization?)

Grading (8pt/10pt)

- (4pt) Submit code to **NTU COOL**
- (1pt) Public simple baseline
- (1pt) Public strong baseline
- (1pt) Private simple baseline
- (1pt) Private strong baseline

Grading -- Kaggle

#	Team Name	Notebook	Team Members	Score
	----- strong baseline -----			0.76023
	----- simple baseline -----			0.68334

Bonus (0.5pt)

- **If you get full marks in this part**, we will make your code **public** to the class.

- In this case, if you also submit **a PDF report briefly describing your methods** (<100 words in English), you get a bonus of **0.5 pt**.
(your report will also be available to all students)

- Report template

Kaggle Submission

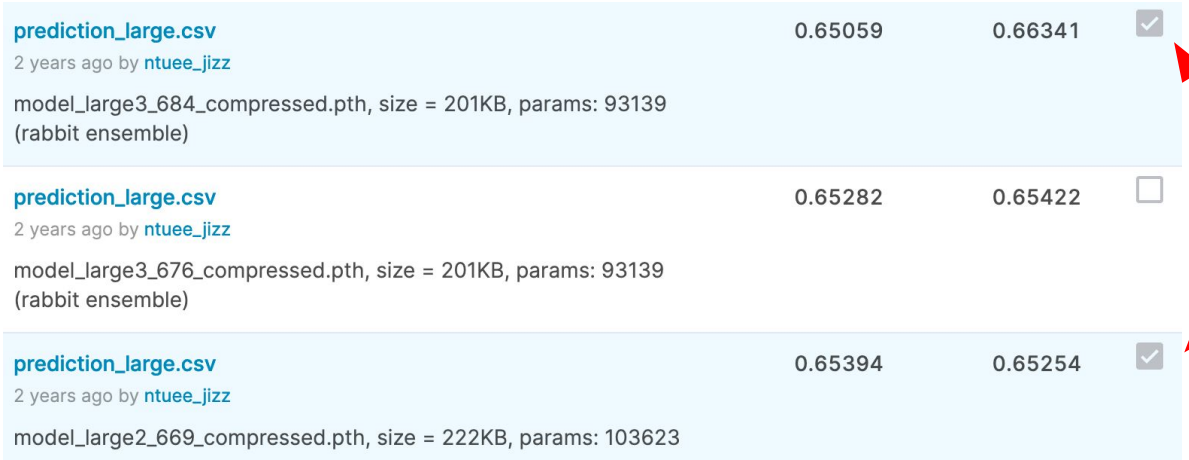
Kaggle Link: https://www.kaggle.com/c/ml2021spring-hw2

- Displayed name: **<student ID>_<anything>**
  - e.g. `b06901020_puipui`
  - For auditing, don't put your student ID in your displayed name.
- Submission format: **`.csv`** file
- Evaluation metric: accuracy
- Submission deadline:
  - **2021/04/02 23:59 (UTC+8)**

```
1 Id,Class
2 0,0
3 1,0
4 2,0
5 3,0
6 4,0
```

Kaggle Submission

- You may submit up to **5** results each day (UTC).
- Up to **2** submissions will be considered for the private leaderboard.



remember to select **2** results for your final scores before the competition ends!

Code Submission

- Compress your code and report, then submit it to NTU COOL.

<student ID>_hw2.zip

e.g. b06901999_hw2.zip

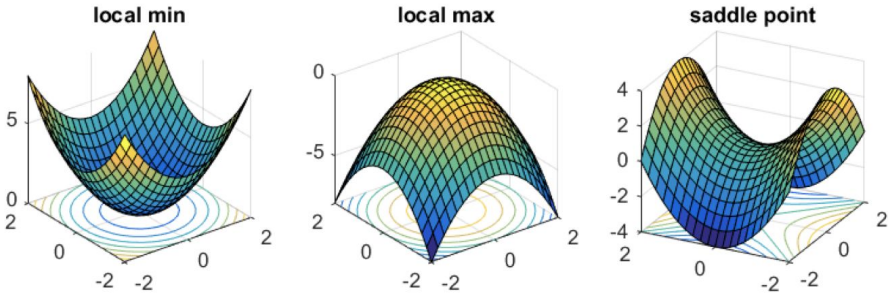
- We can only see your last submission.
- Do not submit your model or dataset.
- If your code is not reasonable, your final grade will be multiplied by 0.9!
- Submission deadline:
    - **2021/04/04 23:59 (UTC+8)**

Code Submission

- Your .zip file should include only
  - **Code**: either .py or .ipynb
  - **Report**: .pdf (only for those who got 8 points in part one)
- Example:

2-2 Hessian Matrix

Task Introduction



Task: Hessian Matrix

Imagine we are training a neural network, and we try to find out whether the model reaches a **local minima-like point, saddle point, or none of the above**. We can make our decision by calculating the Hessian matrix.

What is Hessian?

Hessian is the second order partial derivatives of a model. It is highly recommended to watch the lecture video before starting this part.

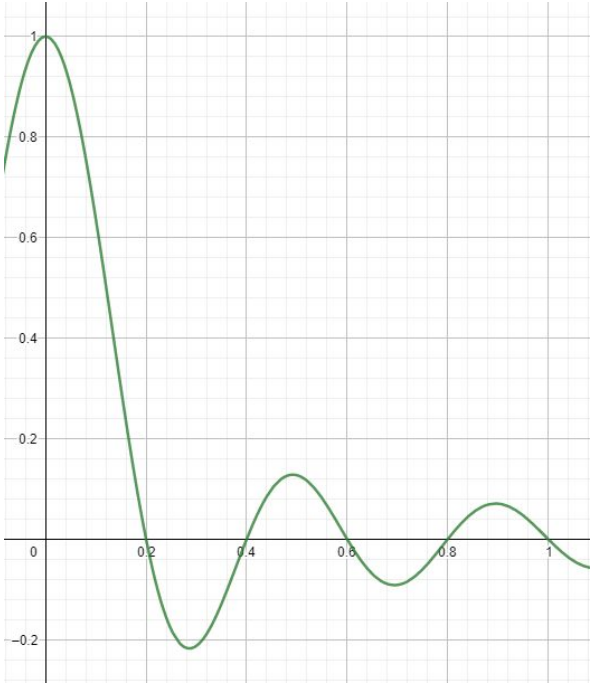
Task Introduction

The target function in this task is a one-variable sinc function.

You will get

- a model checkpoint trained by TA,
- a batch of training data,
- a loss function.

You will calculate the Hessian matrix and make the decision accordingly.



Gradient Norm / Minimum Ratio

1. Gradient Norm

In a normal training process, we rarely have gradients equal to zero. In this homework, we regard those gradient norm less than 1e-3 as zero.

2. Minimum Ratio

For an ideal local minima, **all the eigenvalues** of the hessian matrix are greater than zero. We define the proportion of positive eigenvalues as minimum ratio.

In this homework, if minimum ratio is greater than 0.5 and gradient norm is less than 1e-3, then we assume that the model is at “local minima like”.

$$\text{Minimum ratio} = \frac{\text{Number of } \mathbf{Positive} \text{ Eigen values}}{\text{Number of Eigen values}}$$

Gradient Norm / Minimal Ratio

In this homework, we assume that

- gradient norm < 1e-3 and minimum ratio > 0.5 => **local minima like**,
- gradient norm < 1e-3 and minimum ratio <= 0.5 => **saddle point**,
- gradient norm >= 1e-3 => **none of the above**.

Important Notice

- You don’t need to and shouldn’t change any part of the code.
- You can only use colab to run the code. Otherwise, your result might differ due to environmental issue.
- You will get a different checkpoint according to your student ID, so please make sure to fill in your student ID in the sample code correctly.

```
student_id = 'your_student_id' # fill with your own student ID

assert student_id != 'your_student_id', 'Please fill out your student_id before you start.'
```

Sample Code

Colab Link:
https://colab.research.google.com/github/ga642381/ML2021-Spring/blob/main/HW02/HW02-2.ipynb

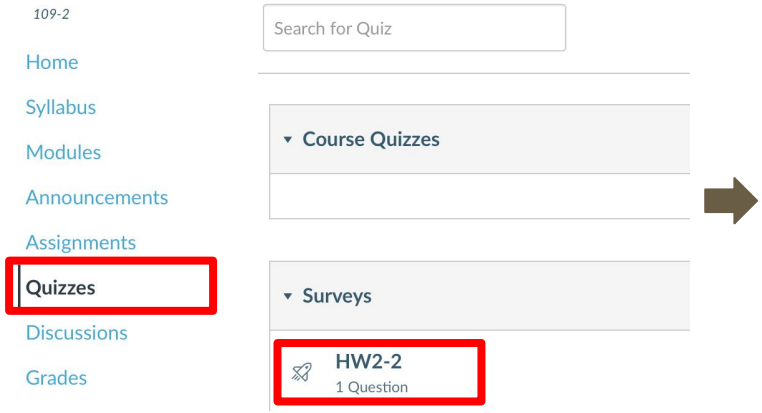
- After executing the sample code, you should get a result like this.
- Notice that each student will get a different answer, so your answer may differ from the example.
- Choose your answer from **local minima like, saddle point, or none of the above**.

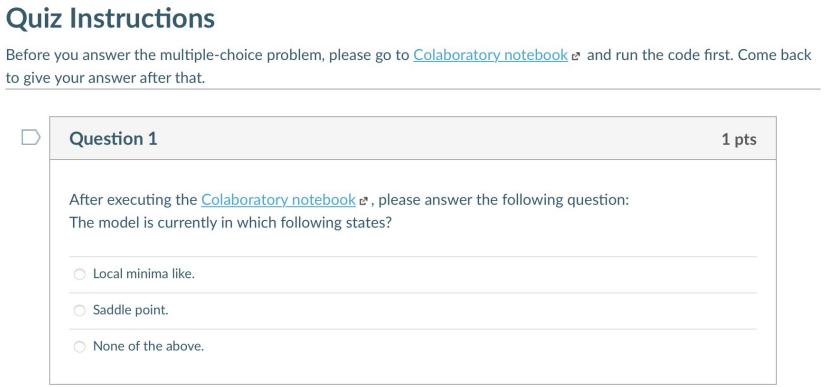
Grading (2pt/10pt)

- (2pt) Correct answer.

NTU COOL Submission

- After you choose your answer, submit it to NTU COOL.
- You can change your answer multiple times before the deadline.
- Submission deadline:
  - **2021/04/02 23:59 (UTC+8)**





Deadlines

- **2-1**
  - Kaggle: **2021/04/02 23:59 (UTC+8)**
  - NTU COOL: **2021/04/04 23:59 (UTC+8)**

- **2-2**
  - NTU COOL: **2021/04/02 23:59 (UTC+8)**

Regulation

- You should NOT plagiarize, if you use any other resource, you should cite it in the reference. (*)
- You should NOT modify your prediction files manually.
- Do NOT share codes or prediction files with any living creatures.
- Do NOT use any approaches to submit your results more than 5 times a day.
- **Do NOT search or use additional data or pre-trained models.**
- Your **final grade x 0.9** if you violate any of the above rules.
- Prof. Lee & TAs preserve the rights to change the rules & grades.

(*) Academic Ethics Guidelines for Researchers by the Ministry of Science and Technology

If any questions, you can ask us via...

- NTU COOL (recommended)
  - https://cool.ntu.edu.tw/courses/4793
- Email
  - ntu-ml-2021spring-ta@googlegroups.com
  - The title should begin with “[hwX]” (X is the homework number)
- TA hour
  - Each Friday during class