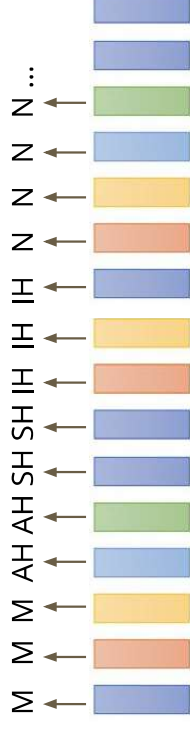


Task Introduction



Task: Multiclass Classification

Frame-wise phoneme prediction from speech.



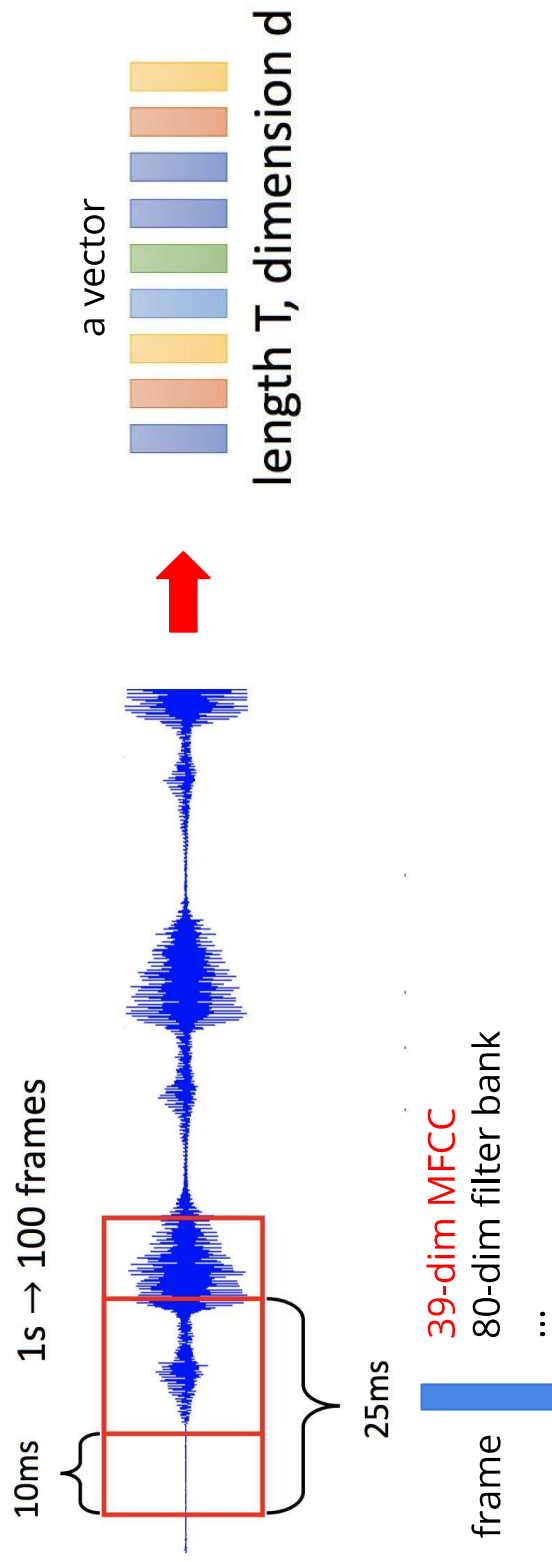
What is a phoneme?

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

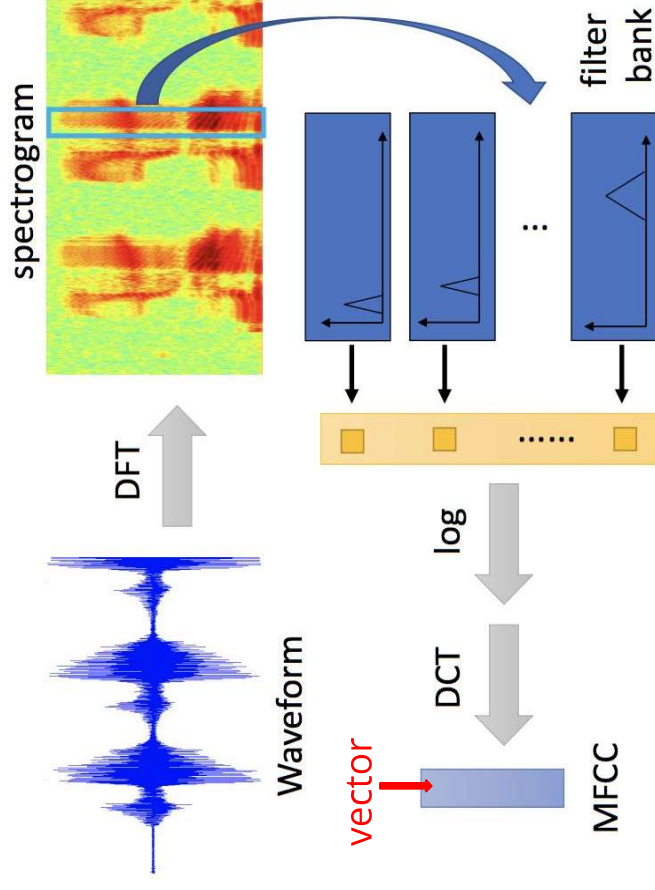
Task Introduction

Data Preprocessing



Task Introduction

Acoustic Features - MFCCs (Mel Frequency Cepstral Coefficients)



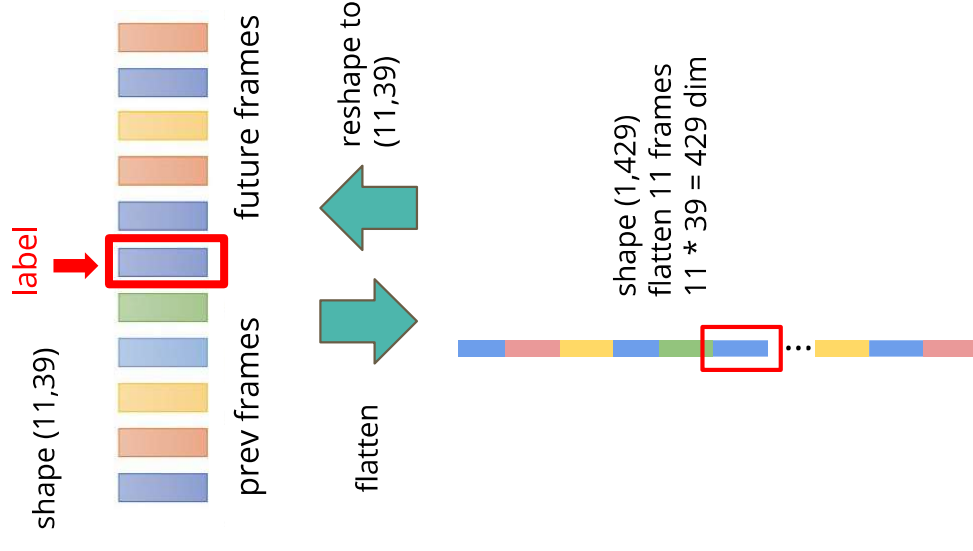
For more details,
please refer to Prof. Lin-Shan Lee's
[\[Introduction to Digital Speech Processing\]](#)
[Chap.7](#)

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!



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 = 11×39)
 - 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	beet	13	l	lay	26	dx	mud dy
1	ih	bit	14	r	ray	27	g	g ay
2	eh	bet	15	y	yacht	28	p	p ea
3	ae	bat	16	w	way	29	t	t ea
4	ah	but	17	er	bird	30	k	k ey
5	uw	boot	18	m	m om	31	z	z one
6	uh	book	19	n	n oon	32	v	v an
7	aa	bob	20	ng	sing	33	f	f in
8	ey	ba it	21	ch	cho ke	34	th	th in
9	ay	bite	22	jh	j oke	35	s	s ea
10	oy	boy	23	dh	th en	36	sh	sh e
11	aw	bou t	24	b	b ee	37	hh	h ay
12	ow	boa t	25	d	d ay	38	sil	silence/closure sounds

