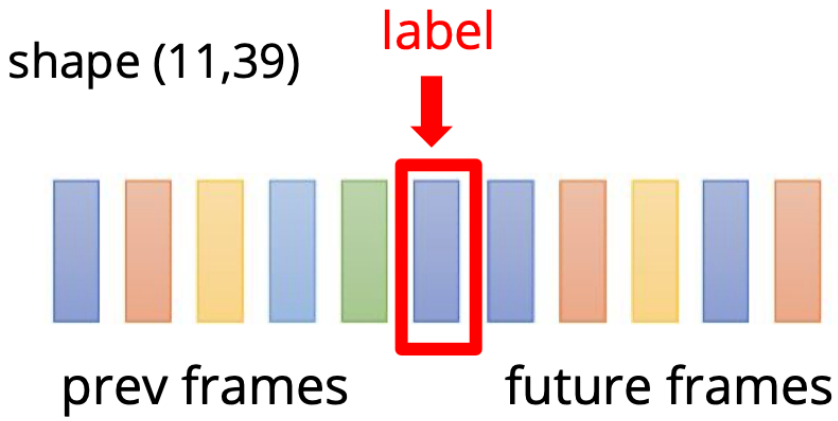
2023年6月14日 星期三

Homework 2

Phoneme Classification

- Description:

  將一段語音訊號輸入進行 Phoneme Classification，其中每段為長度39的向量，並給予前後五段的數據，因此一個特徵為 11 * 39 = 429 維度，如下圖。



  共 39 組類別如下圖。

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***oy***	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

- Architecture：

    簡單的 Linear + LeakyReLU + BN + Dropout 模型

```
import torch
import torch.nn as nn

class Classifier(nn.Module):
    def __init__(self):
        super(Classifier, self).__init__()
        self.net = nn.Sequential(
            nn.Linear(429, 2048),
            nn.LeakyReLU(),
            #nn.ReLU(),
            nn.BatchNorm1d(2048),
            nn.Dropout(0.45),
            nn.Linear(2048, 2048),
            nn.LeakyReLU(),
            #nn.ReLU(),
            nn.BatchNorm1d(2048),
            nn.Dropout(0.45),
            nn.Linear(2048,1024),
            nn.LeakyReLU(),
```

- HyperParameter：

	Value
BATCH_SIZE	2048
VAL_RATIO	0.01
learning_rate	0.001
num_epoch	64

- Criterion：CrossEntropyLoss ( )
- Optimizer：NAdam
- Standard:

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

- Result:
  - Private Score:0.75176
  - Public Score:0.75246