

專題研究

WEEK 5 – Deep Neural Networks in Kaldi

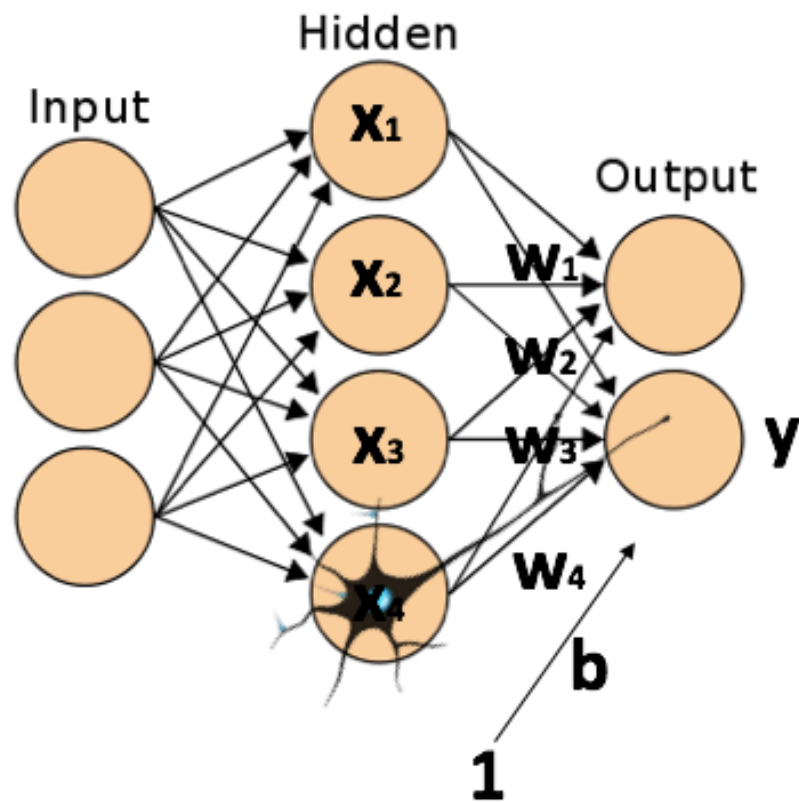
Prof. Lin-shan Lee

TA. 鍾佩宏 (r05942048@ntu.edu.tw)

Outline

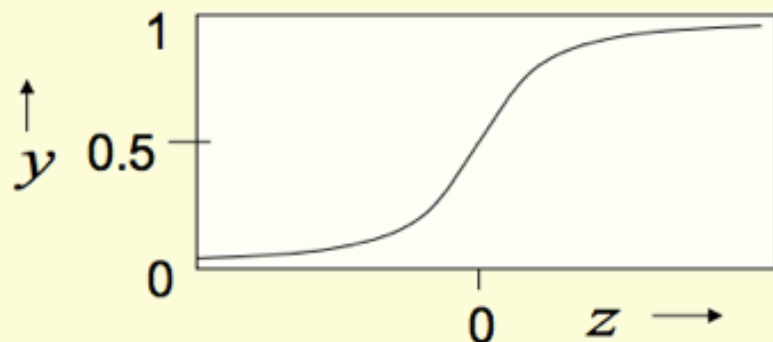
- **Recap**
- **Homework**
- **Script**
- **TODO**
- **Q&A**

Neural Network



$$y_k = f\left(b_k + \sum_i x_i w_{ik}\right)$$

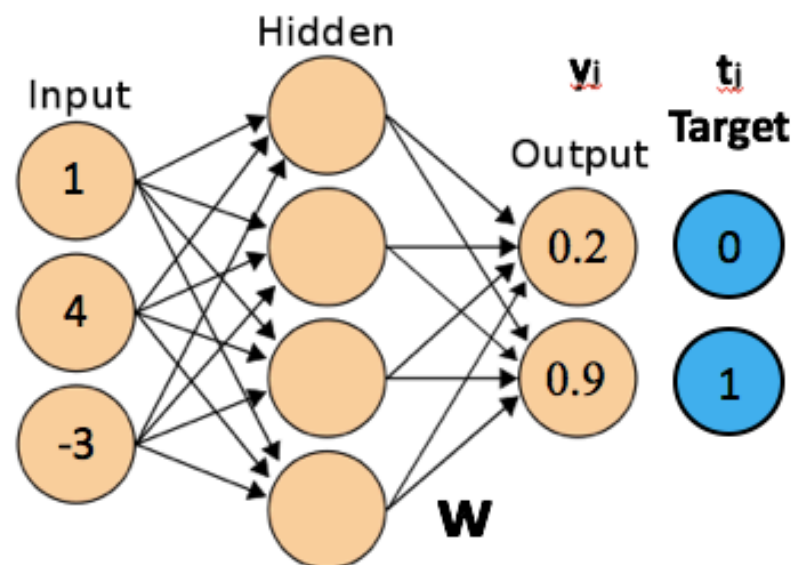
$$z = b + \sum_i x_i w_i \quad y = \frac{1}{1 + e^{-z}}$$



- A lot of simple non-linearity \rightarrow complex non-linearity

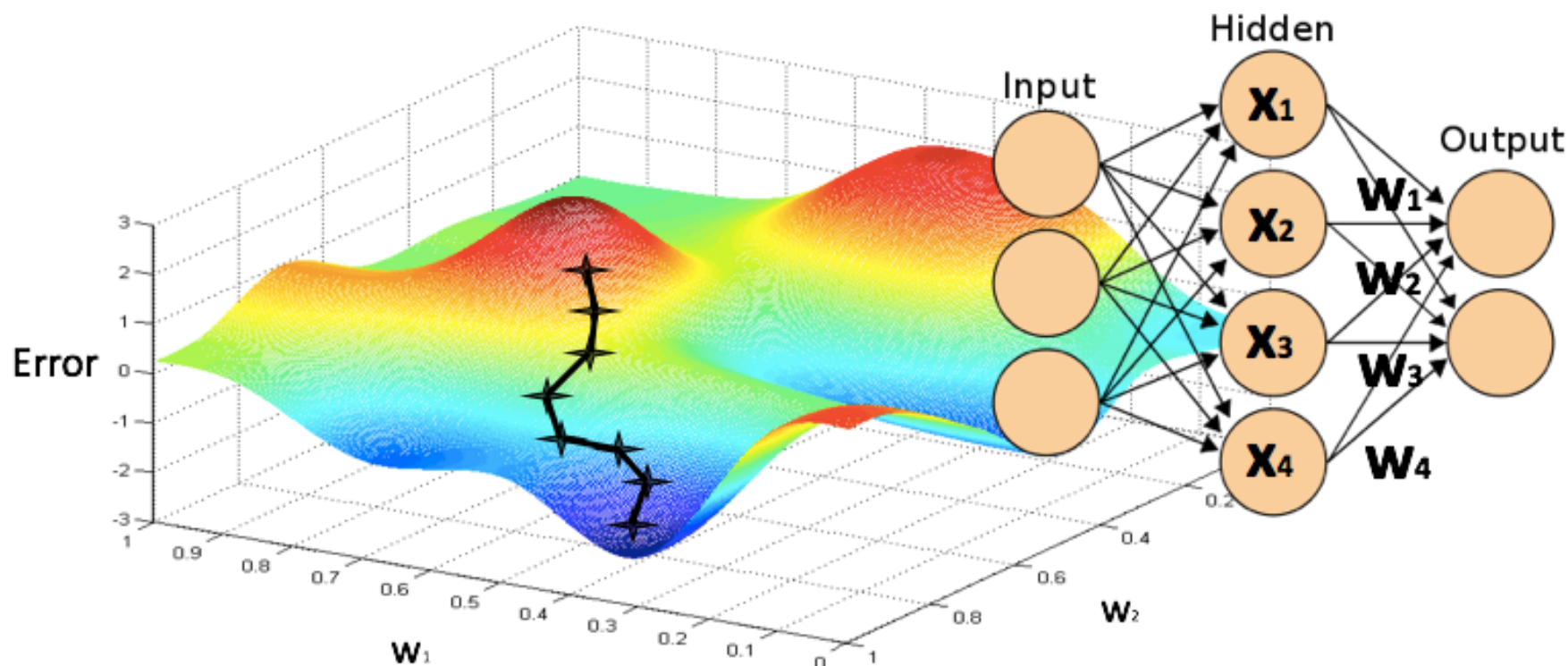
Neural Network Training – Back Propagation

- **Start with random weights**
- **Compare the outputs of the net to the targets**
- **Try to adjust the weights to minimize the error**



$$E = \frac{1}{2} \sum_{j \in \text{output}} (t_j - y_j)^2$$

Gradient Descent Algorithm



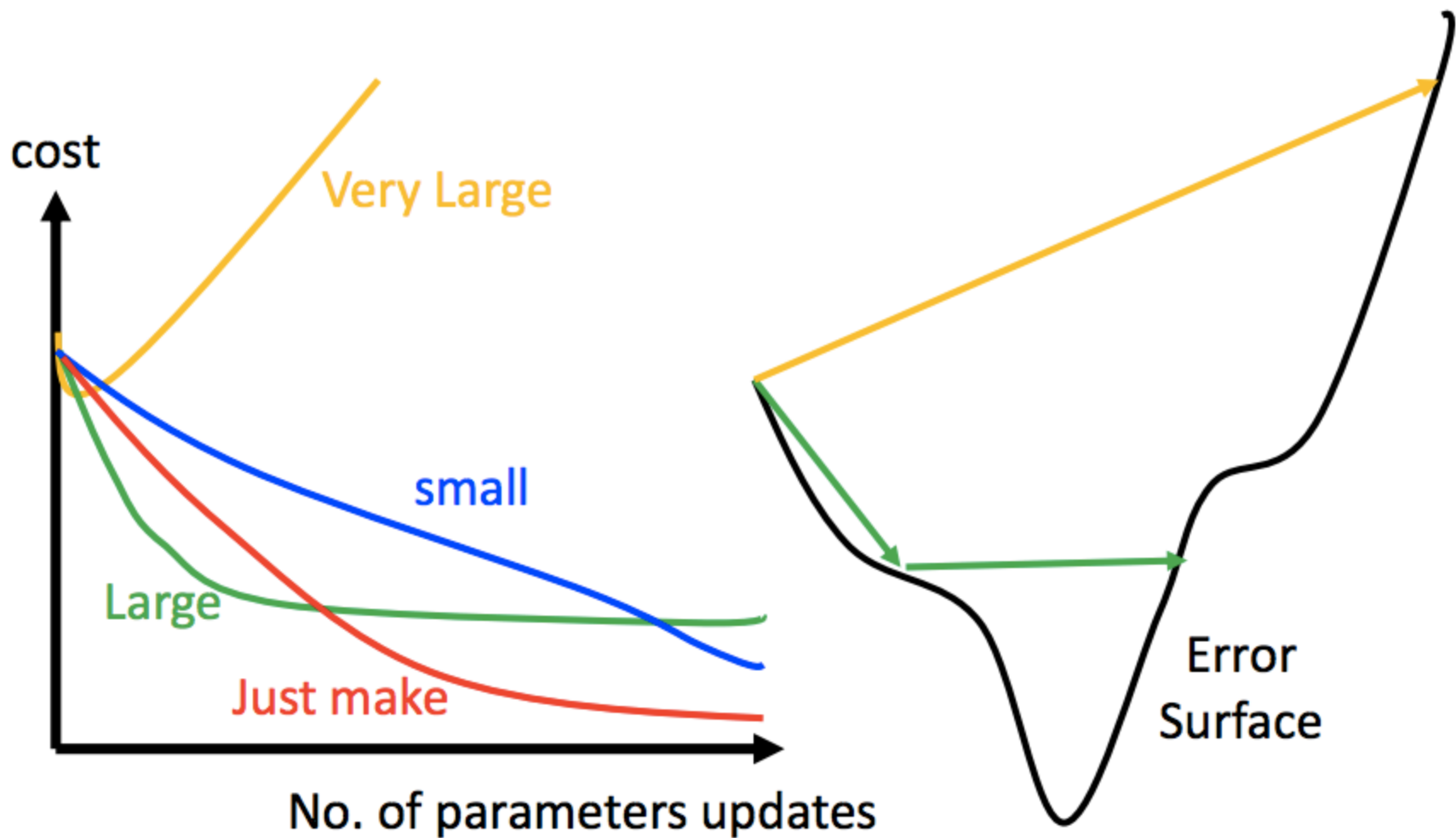
$$w_{t+1} = w_t - \alpha \frac{\partial E}{\partial w}$$

Updated weights \rightarrow w_{t+1} w_t \rightarrow Learning rate α
Weight at t -th iteration

Learning Rate

$$\theta^i = \theta^{i-1} - \eta \nabla C(\theta^{i-1})$$

- Set the learning rate η carefully



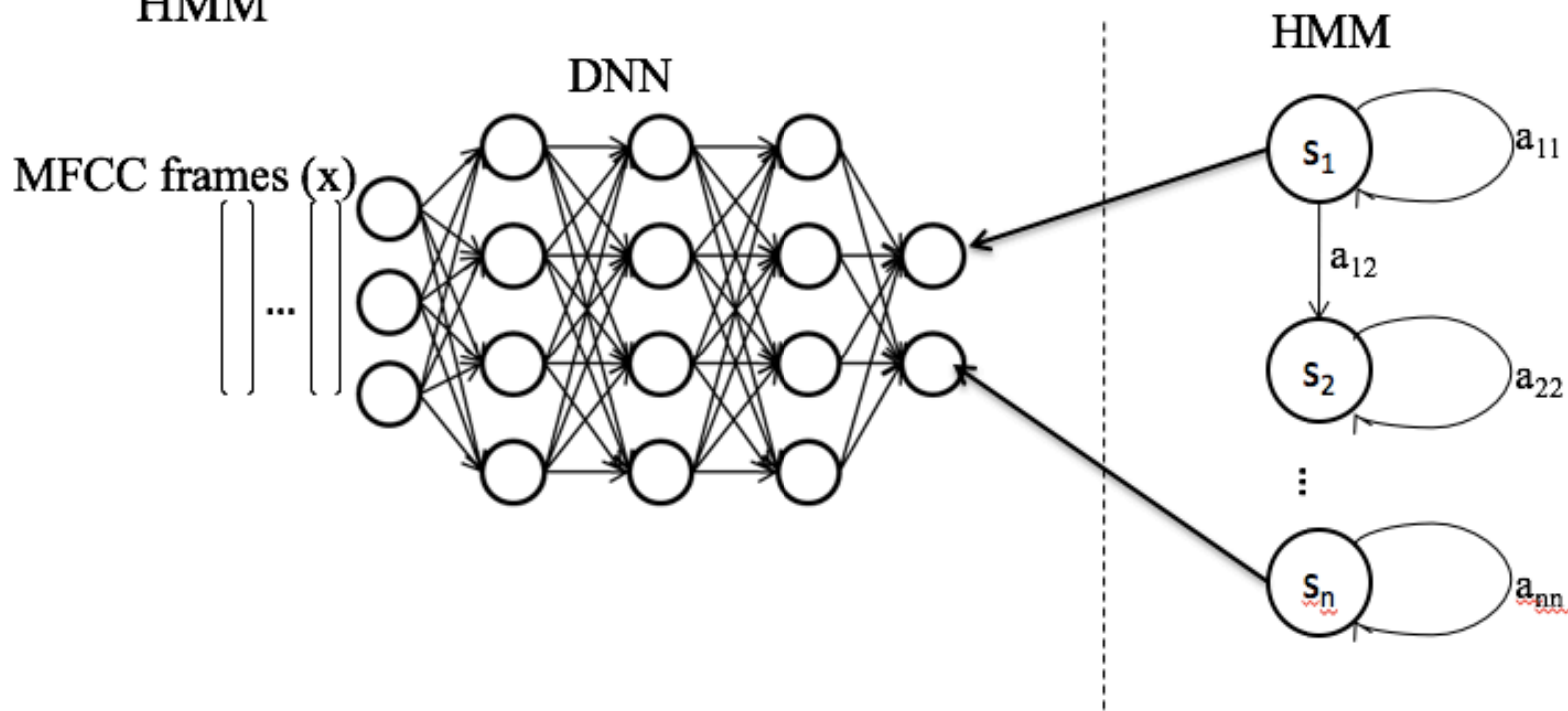
Deep Neural Network for Acoustic Modeling

- **DNN as triphone state classifier**

- input: acoustic features, e.g. MFCC
- output layer of DNN representing triphone states
- fine tuning the DNN by back propagation using labelled data

- **Hybrid System**

- normalized output of DNN as posterior of states $p(s|x)$
- state transition remaining unchanged, modeled by transition probabilities of HMM



Homework

- 08.mlp.train.sh
- 09.mlp.decode.sh
- 閱讀：數位語音處理導論 ch9

Script

- 登入工作站 (pietty/putty/Xshell等terminal)
ssh 140.112.21.80 port 22
 - 複製壓縮檔到自己的子資料夾
cp /share/week5.tar.gz
 - 解壓縮
tar -zxvf week5.tar.gz
- ```
chung95191@SpeechLab531:~/proj1.ASTMIC.subset/week5$ ls
08.mlp.train.sh 09.mlp.decode.sh align_dev.sh setup.sh
```
- cp setup.sh /proj1.ASTMIC.subset
  - cp 08.mlp.train.sh /proj1.ASTMIC.subset/script
  - cp 09.mlp.train.sh /proj1.ASTMIC.subset/script

# To Do

- ❑ **Step 1: Execute the following command:**  
bash setup.sh  
script/08.mlp.train.sh  
script/09.mlp.decode.sh
- ❑ **Observe the accuracy**
- ❑ **Step 2.1: Tune the NN model parameter in script/08.mlp.train.sh**
- ❑ depth
- ❑ num\_hidden
- ❑ minibatch\_size

```
30 #NN parameter part
31 depth=2
32 rbm_depth=0
33 rbm_hidd_num=1024
34 num_hidden=1024
35 minibatch_size=256
```

# To Do

- **Step 2.2: Modify the training parameters in script/08.mlp.train.sh**
- **learning rate**

```
39 #training parameters
40 max_iters=50
41 min_iters=15
42 keep_lt_iters=10
43 learn_rate=0.1
44 start_halving_impr=0.01
45 end_halving_impr=0.001
46 halving_factor=0.5
47
```

- **Step 2.3: Tune the feature context parameter**
- **context**

```
48 #feature part
49 feat_train=$train_feat_mlp
50 feat_dev=$dev_feat_mlp
51 feat_test=$test_feat_mlp
52 context=0
53
```

# To Do

- Step 3: Tune the acoustic weights in 09.mlp.decode.sh

```
50 echo "=====Finding the optimal acoustic weight====="
51 for x in `seq 9 10`; do
52 acwt=`perl -e "printf (\"%4.2f\", $x * 0.02);"`
53 log=$dir/log/dev.${acwt}.log
54 echo "=====
55 echo " Generating char level transcription on acoustic weig
56 echo " Then evaluate the accuracy."
57 echo " log -> $log"
58 echo "=====
```

- acoustic weight
- Step 4: Repeat Step2 and Step3 to get the best accuracy

# Q&A

- **If you got any question, contact TA through FB group or email instantly.**
  - **Feel free saying about anything.**
  - **Email:**  
**r05942048@ntu.edu.tw**  
**鍾佩宏**