

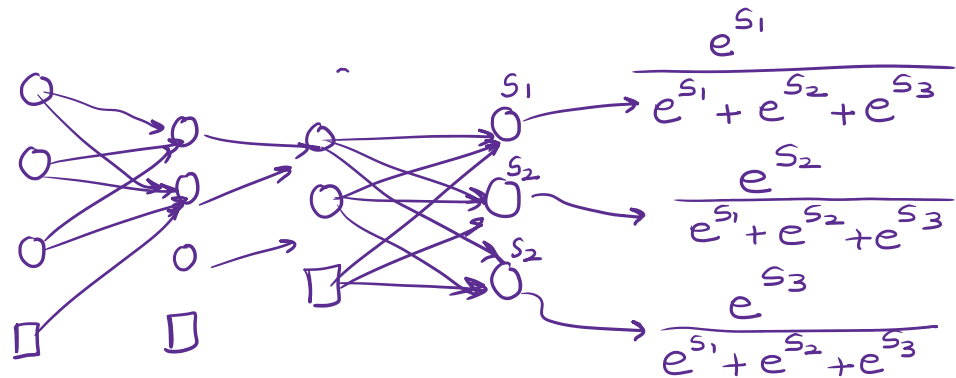
$$f(x_i) = \frac{e^{x_i}}{\sum_{j=1}^m e^{x_j}}; m \text{ being the number of different categories}$$

Softmax
function

Multi-class classification

$y: A \text{ or } B \text{ or } C$

- Used for multi-class classification problems
- It is a generalization of sigmoid function
- It produces output between 0 and 1

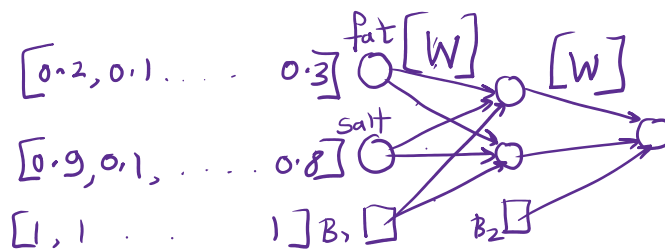


Online training

Obs.	Fat	Salt	Acceptance
1	0.2	0.9	like
2	0.1	0.1	dislike
3	0.2	0.4	dislike
4	0.2	0.5	dislike
5	0.4	0.5	like
6	0.3	0.8	like

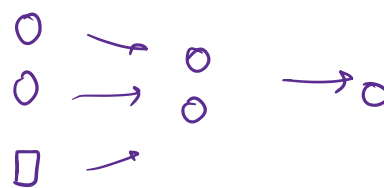
Batch training

Obs.	Fat	Salt	Acceptance
1	0.2	0.9	like
2	0.1	0.1	dislike
3	0.2	0.4	dislike
4	0.2	0.5	dislike
5	0.4	0.5	like
6	0.3	0.8	like

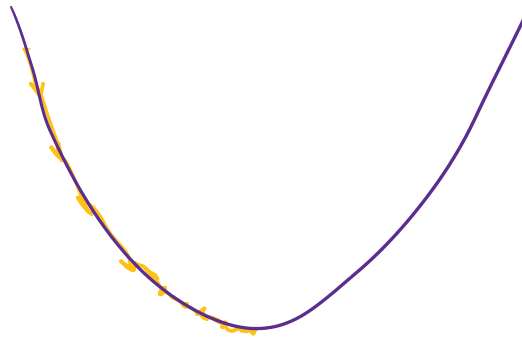
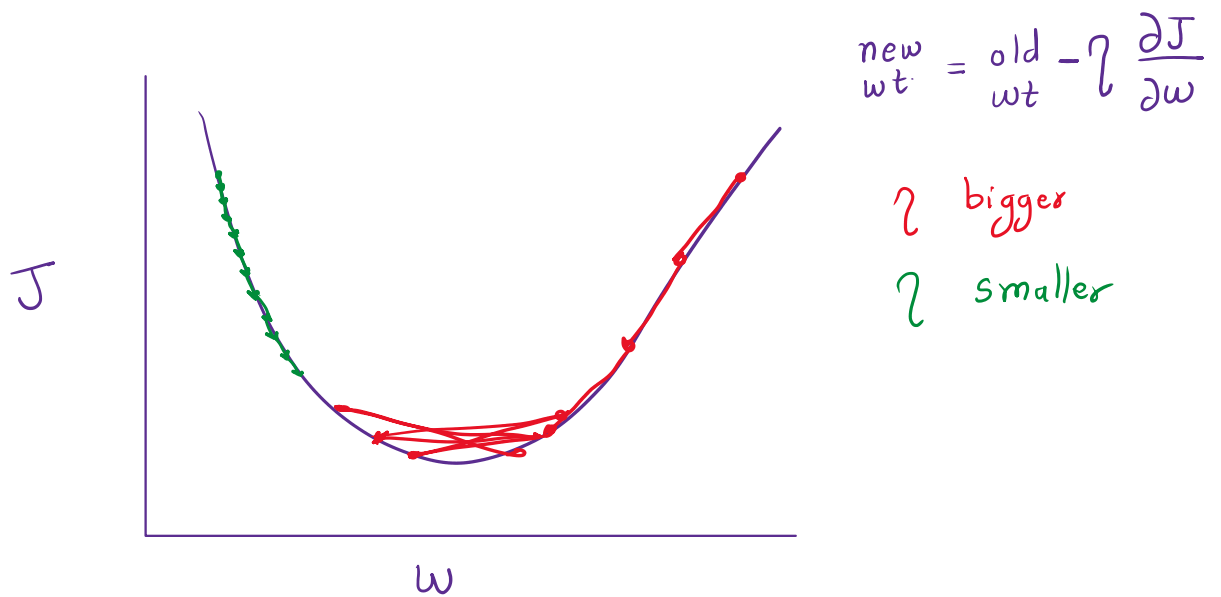


Batch-Size Training

↓ Lakh
 ↳ 10000 }
 10000 } 10
 ... }
 10000 } batches



Keep batch size small, then mini-batch-size training



Min max Scaling : $\frac{X - \min(X)}{\max(X) - \min(X)}$

$\min(X) = 2$
 $\max(X) = 13$

X

5 $(5-2)/(13-2)$

8 $(8-2)/(13-2)$

2 $(2-2)/(13-2) = 0$

9 $(9-2)/(13-2)$

13 $(13-2)/(13-2) = 1$

(20,10,5)

24
0
...

20
0
...

10
0
...

5
0
...

0

□

□

□

□

Regularizers

$$l_1 := \sum_i |\beta_i|$$

$$l_2 := \sum_i \beta_i^2$$

$$l_1 - l_2$$