Activation function:

- Sigmoid
- ReLU
- Softmax
- tanh

Softmax: → multiclass classification

Output layer

z_1 (1.5) $\quad e^{1.5} \rightarrow \text{Softmax}(z_1) = \dfrac{e^{1.5}}{\Sigma} = \dfrac{4.48}{7.24} = 62\%$

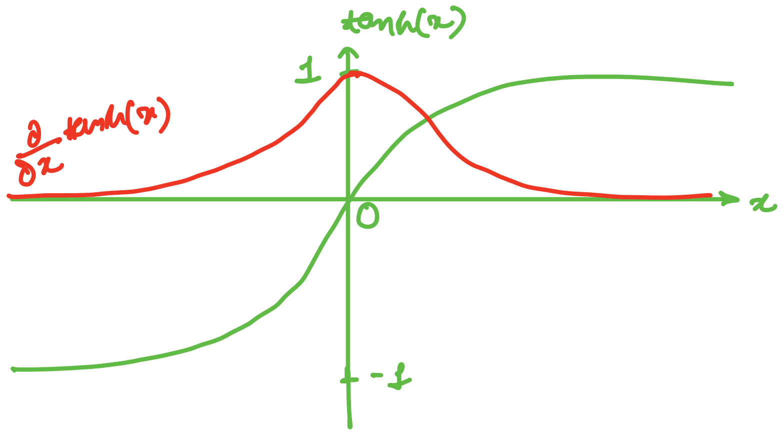
z_2 (0.5) $\quad e^{0.5} \rightarrow \text{Softmax}(z_2) = \dfrac{e^{0.5}}{\Sigma} = 23\%$

z_3 (0.1) $\quad e^{0.1} \rightarrow \dfrac{e^{0.1}}{\Sigma} = 15\%$

$$\Sigma = e^{1.5} + e^{0.5} + e^{0.1} = 7.24$$

Tanh

$$f(x) = \tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} = \frac{\sinh(x)}{\cosh(x)}$$



$$\frac{\partial}{\partial x} \tanh(x) = 1 - \tanh^2(x)$$

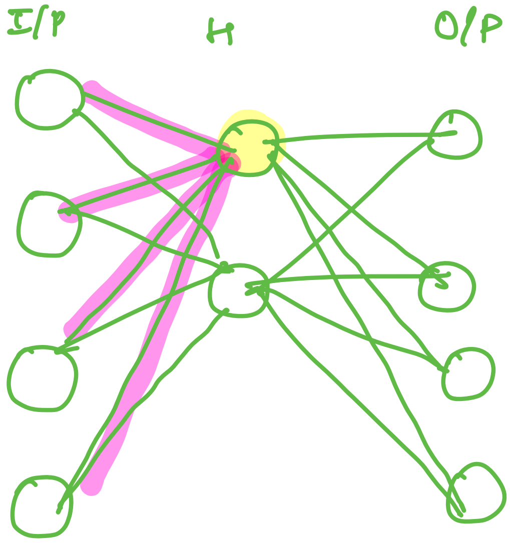
Overfitting in NN:

- Training Data size large.

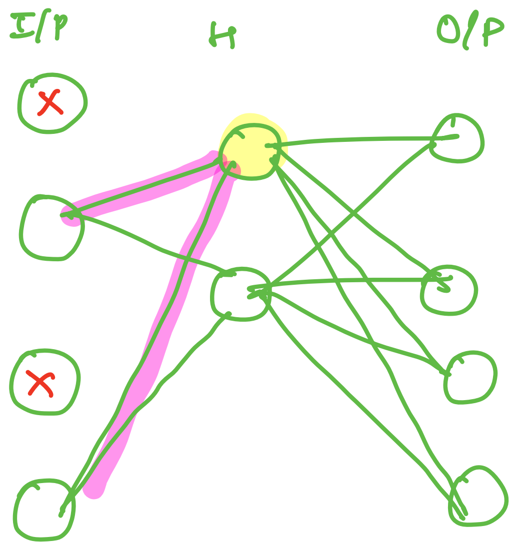
- Model complexity less

- Dropout

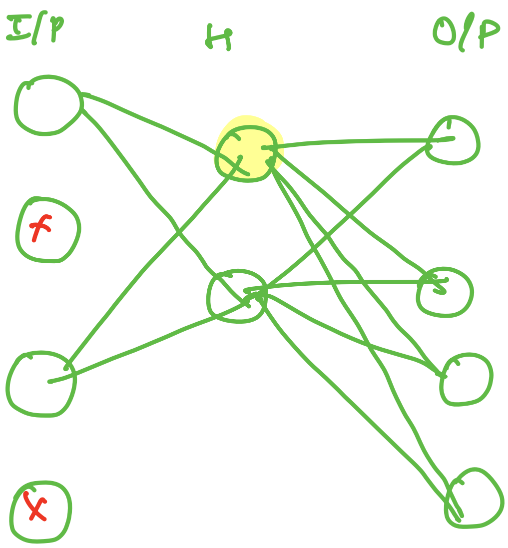
0.5 layer



epoch 1



epoch 1



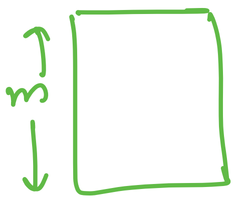
- Regularisation

$$L = \frac{1}{2} \sum (y-\hat{y})^2 + \frac{\lambda}{2} \|W\|_2 \rightarrow w_1^2 + w_2^2 + w_3^2 \cdots w_n^2$$

Gradient descent Variants

$$w_1 = w_1 - \eta \frac{\partial L}{\partial w_1}$$

① Batch gradient descent



- epochs loop run
  - loss calculate

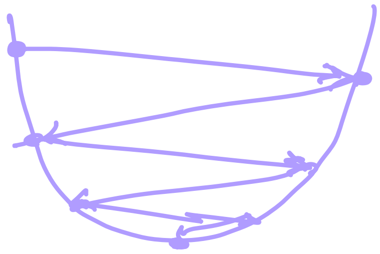
$$L = \frac{1}{2m}\sum_{i=1}^{m}(\hat{y} - y)^2$$

Images:
$10^5 \times 10MB$
10^6 MB

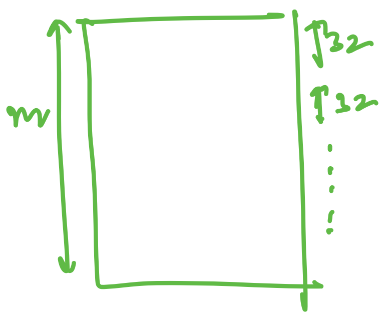
② Stochastic gradient descent

at a time loss calculate
wrt 1 example

- epochs run loop
  - m examples
    - $L = (\hat{y} - y)^2$
    - ∂L
    - w update



③ Mini Batch Gradient Descent



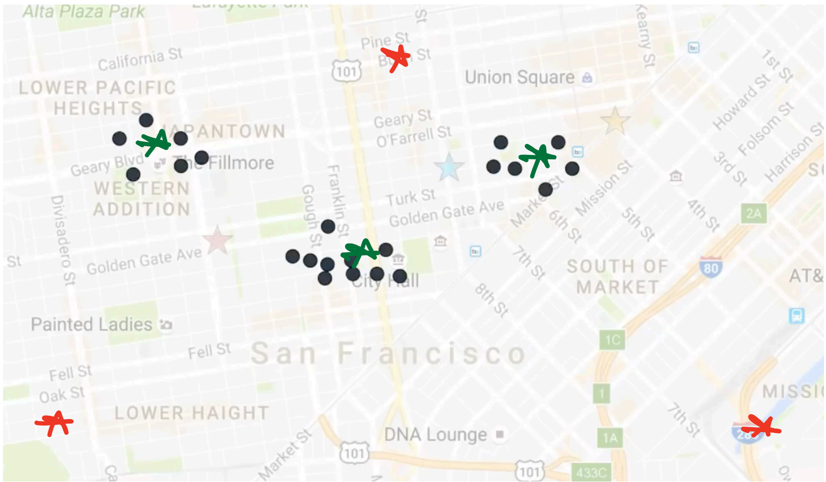
BS = 32
no of batches = $\frac{m}{32}$

- epochs loop
  - no of batches → m/32
    - loss calculate wrt 32 examples

Supervised L.A

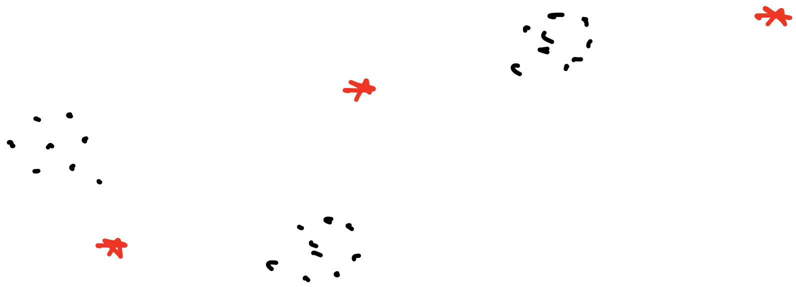
- Linear Reg.
- Log. Reg.
- NB
- DT
- KNN
- ANN

Unsupervised LA : Clustering : group similar kind of data together.
↳ no y label

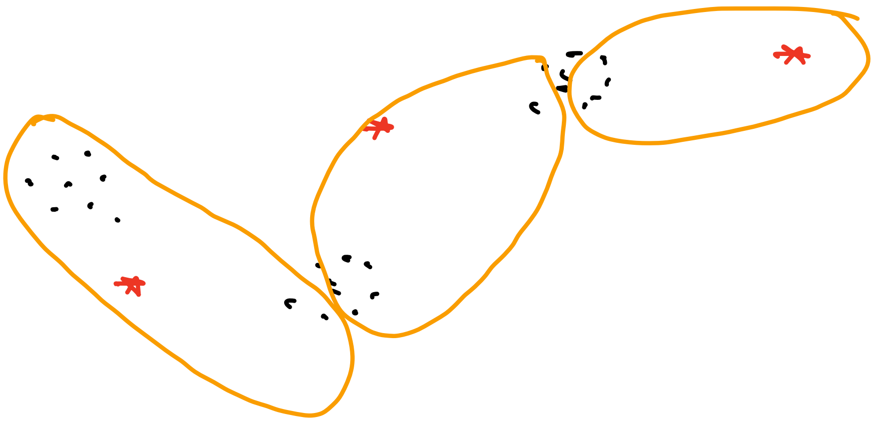


K Means

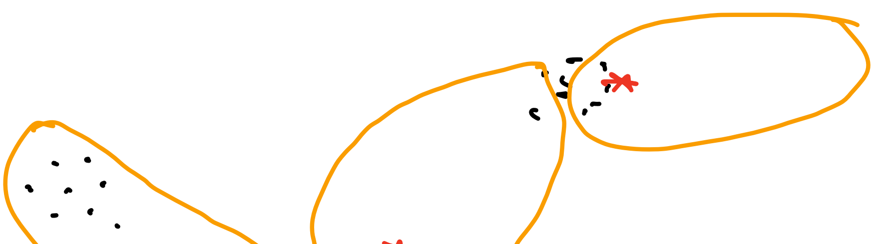
1. Randomly initialize 3 center points

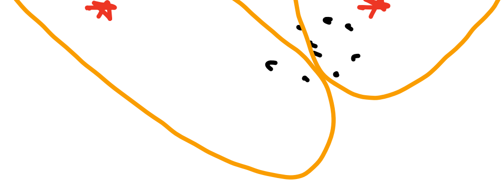


2. Assign each customer to its nearest center point



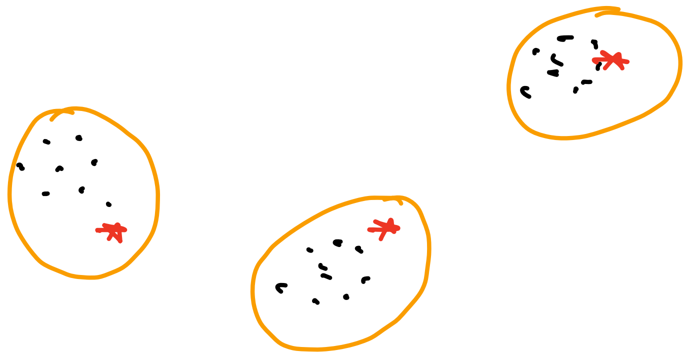
3. Update the center location by taking the mean of points assigned to the cluster.



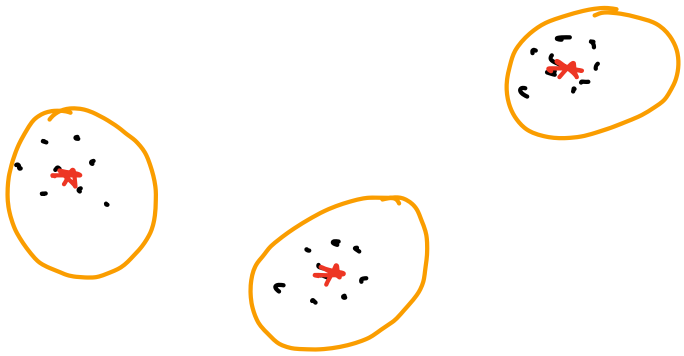


Step 2 & 3 repeat

Step2:



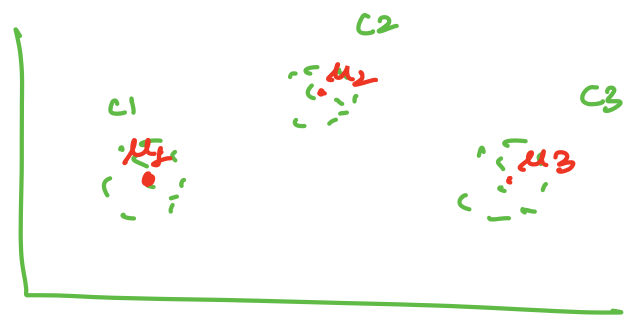
Step3:



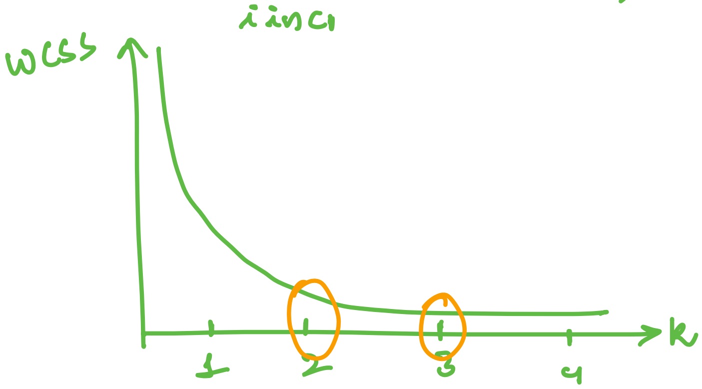
Elbow Method: to find k in kmeans?

k = 2
3
4
5
6
7
⋮

WCSS (within Cluster Sum of Squares)



$$WCSS = \sum_{i \in C_1} distance(x^i, \mu_1) + \sum_{i \in C_2} distance(x^i, \mu_2) + \dots$$



A meteorological research team is analyzing a simplified dataset containing readings from different weather stations. The dataset is given in Table IV with features: Temperature and Humidity. The team decides to use the k-means clustering algorithm to categorize these samples into two distinct groups (k = 2) based on their similarities in temperature and humidity. The initial centroids are chosen as (10, 80) for Centroid 1 and (20, 60) for Centroid 2. Perform k-means clustering for two iterations and provide the cluster assignments and centroids after each iteration. Using graphical representations, demonstrate how the intra-cluster distances change across iterations. **Hint:** intra-cluster distance in average distance of all points from centroid in a cluster. [4] [CO2]

Table IV

Sample No.	1	2	3	4	5	6
Temperature (°C)	10	11	10	20	21	20
Humidity (%)	80	79	81	60	59	61

	C_1 Distance (10,80)	C_2 Distance (20,60)	Cluster
10,80	$\sqrt{(10-10)^2+(10-80)^2} = 0$	22.36	C1
11,79	$\sqrt{(11-10)^2+(79-80)^2} = 1.41$	21.02	C1
10,81	1.0	23.26	C1
20,60	22.36	0	C2
21,59	23.71	1.41	C2
20,61	21.47	1	C2

Update center points:

C_1': 10,80 ; 11,79 ; 10,81 $\Big\}$ $\dfrac{10+11+10}{3}, \dfrac{80+79+81}{3}$

$(10.33, 80)$

C_2': 20,60 ; 21,59 ; 20,61 $\Big\}$ $\dfrac{20+21+20}{3}, \dfrac{60+59+61}{3}$

$(20.33, 60)$

	C_1' Distance (10.33, 80)	C_2' Distance (20.33, 60)	Cluster
10,80	0.33	22.51	C_1'
11,79	1.20	21.17	C_1'
10,81	1.05	22.40	C_1'
20,60	22.22	0.33	C_2'
21,59	23.56	1.20	C_2'
20,61	21.32	1.05	C_2'

update center points:

C_1'': 10,80 ; 11,79 ; 10,81 $\Big\}$ $\frac{10+11+10}{3}, \frac{80+79+81}{3}$

(10.33, 80)

C_2'': 20,60 ; 21,59 ; 20,61 $\Big\}$ $\frac{20+21+20}{3}, \frac{60+59+61}{3}$

(20.33, 60)

Iteration 1:

C1 intra cluster distance $= \frac{0+1.41+1.0}{3} = 0.803$

C2 intra cluster distance $= \frac{0+1.41+1.0}{3} = 0.803$

Iteration 2:

C1 intra cluster distance $= \frac{0.33+1.20+1.05}{3} = 0.86$

C2 intra cluster distance $= 0.86$