Date :

Deep Learning CS583 A
Fall 2021
Quiz 2

November 30, 2021

Instructor : Jia Xu

Student Name : Shreeya Kokate

Student ID : 20005256

Student email : skokate@stevens.edu ;
shreeya.csai@gmail.com
address

Date :

1) Difference between supervised & unsupervised learning.

Supervised learning	Unsupervised learning
* These algorithms are trained using labeled data	* These algorithms are trained using unlabeled data
* These model takes direct feedback to check if it is predicting correct output or not.	* These models does not take any feedback
* It predicts the output	* It finds the hidden patterns in data.
* Here, input data is provided to the model along with the output	* Here, only input data is provided.
* The goal is to train the model so it can predict the output	* The goal is to find patterns & useful insights

Date :

when it is given new data.	from the unknown dataset.
* It needs supervision to train the model.	* It doesn't need any supervision.
* Categorized into Classification & Regression problems	* Categorized into Clustering & Association problems.

Date :

Formula for the loss function used in multiclass classification problem

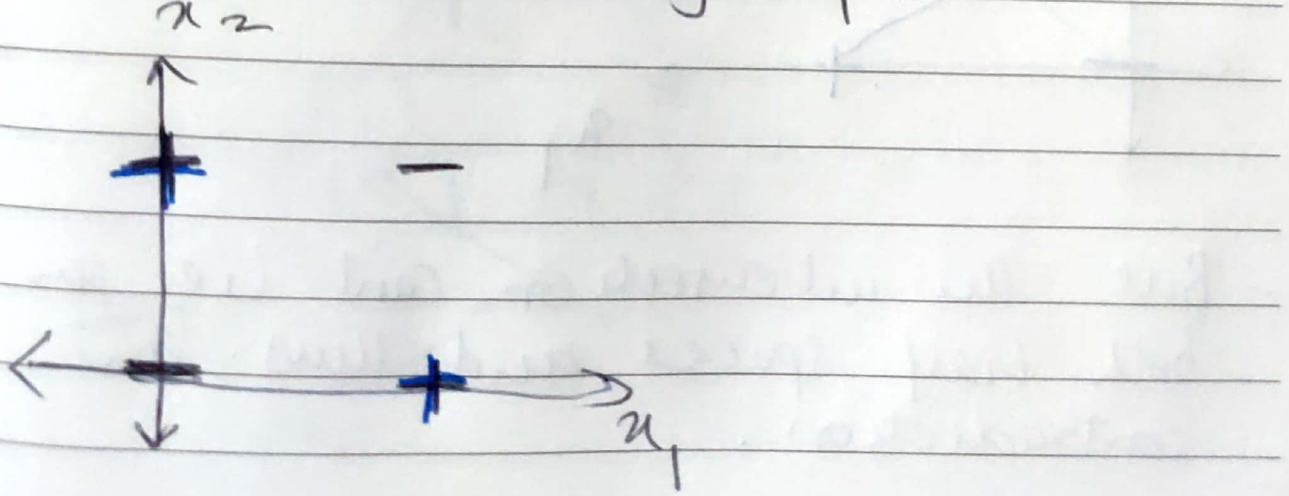
⇒

$$L(\theta) = -\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{k} [y_{ij} \log (p_{ij})]$$

3. XOR

⇒

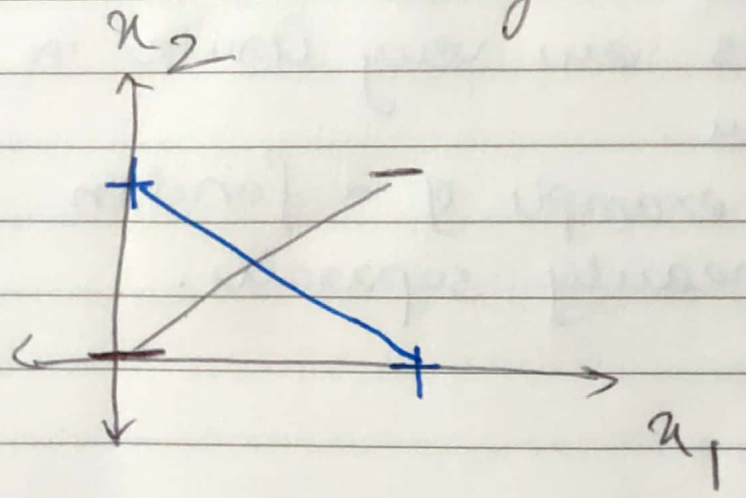
- Single neurons are very limited in expressive power.
- XOR is an example of a function that is not linearly separable.



Date :

x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

- Half-spaces are obviously convex.
- Suppose there were some feasible hypothesis. If the positive examples are in the positive half space, then the blue green line segment must be as well.
- Similarly, the ~~red~~ black line segment within the negative half space



- But the intersection can't lie in both half spaces and thus an contradiction.

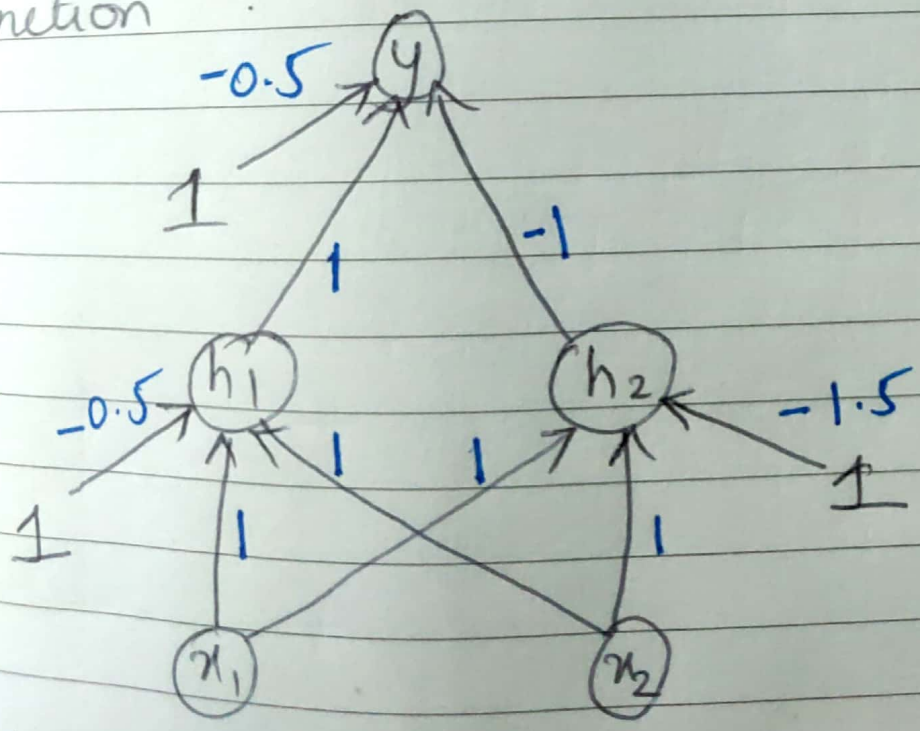
Date :

We can overcome this limitation using feature maps, just like linear regression

x_1	x_2	$\phi_1(x)$	$\phi_2(x)$	$\phi_3(x)$	t
0	0	0	0	0	0
0	1	0	1	0	1
0	0	1	0	0	1
1	1	1	1	1	0

This is linearly seperable.
We now use neural nets to learn non linear hypothesis directly.

XOR Network using threshold activation function :



Date :

4)

⇒ No, it is not possible to construct a single layer neural network with threshold activation function implementing addition of 2 bits because binary addition with a single layer cannot be performed.

5)