# Assignment3

NAME：董骏博

SID：12432995

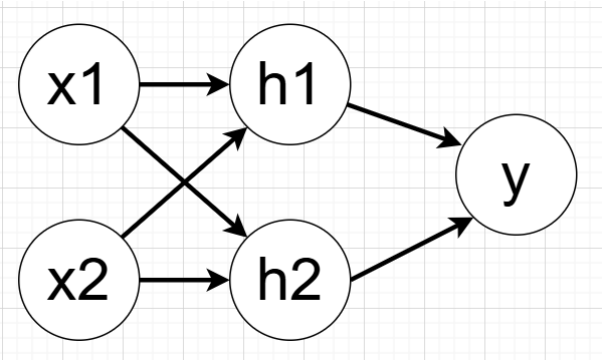
# Task 1

## network structure

- **Input layer:** 2 nodes, input `x_1` and `x_2`
- **Hidden layer:** one layer, 2 nodes
- **Output layer:** 1 nodes,  output `y_hat`
- **Activate function:** `Sigmoid`

  The XOR problem is a binary classification problem, and `Log loss` is applicable to classification problems. Its penalty for wrong predictions is asymmetric. When $y = 1$, the smaller the predicted probability $\hat{y}$, the greater the penalty; when $y = 0$, the larger the predicted probability $\hat{y}$, the greater the penalty.

The network structure is:




**loss function:** `Log-likelihood Loss`

$$L = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log \hat{y}_i + (1-y_i) \log \left(1-\hat{y}_i\right)$$

## The way ANN predict True labels

**Initialization:** Randomly initialize weights and biases.

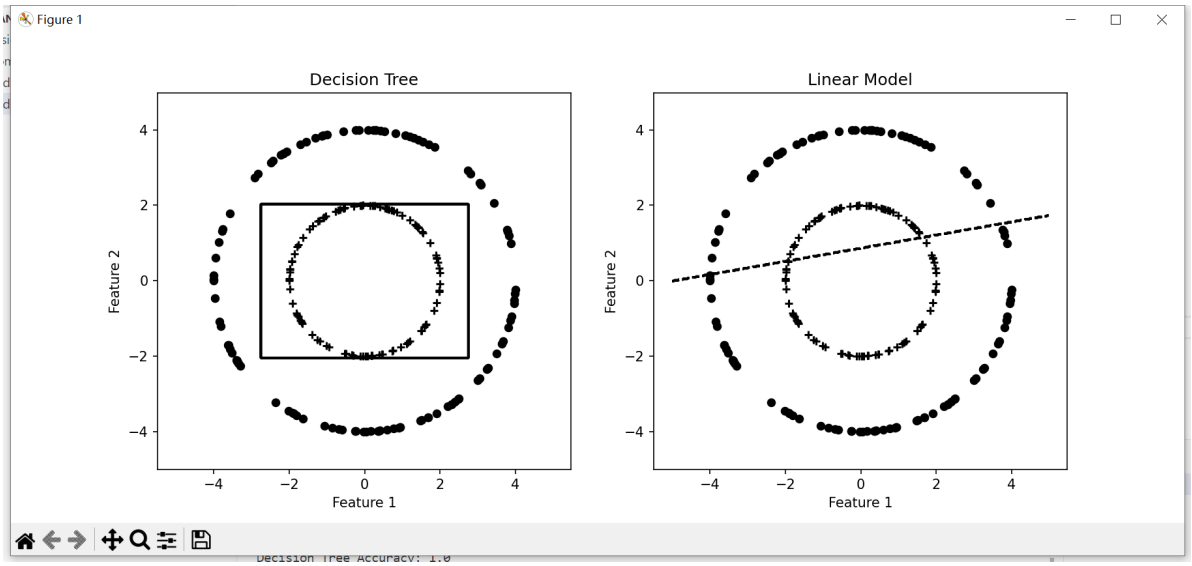
**Forward propagation:** Calculate the output value of each layer of nodes to get the final output of the network.

**Calculate loss:** Use log loss to calculate the error between the predicted result and the actual result.

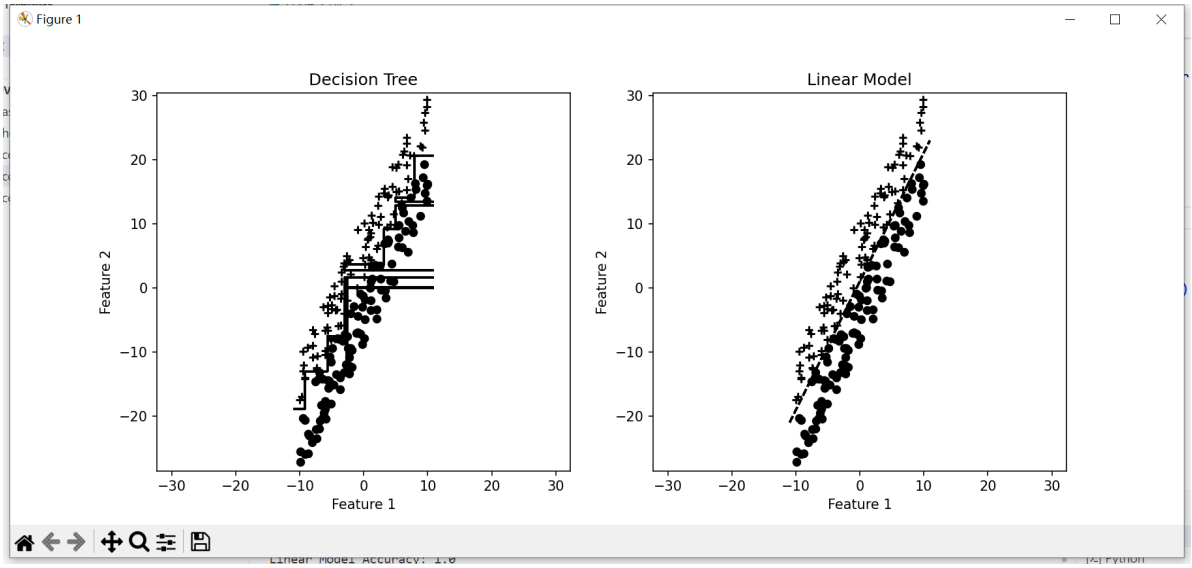
**Backward propagation:** Adjust weights and biases according to gradients and learning rates

# Task 2

### (1)




### (2)




# Task 3

### (1)

$$
\begin{aligned}
E &= C_5^3 \times \epsilon^3 \times (1-\epsilon)^2 + C_5^4 \times \epsilon^4 \times (1-\epsilon)^1 + C_5^5 \times \epsilon^5 \times (1-\epsilon)^0 \\
&= \frac{5!}{3!(5-3)!} \times \epsilon^3 \times (1-\epsilon)^2 + \frac{5!}{4!(5-4)!} \times \epsilon^4 \times (1-\epsilon)^1 + \frac{5!}{5!(5-5)!} \times \epsilon^5 \\
&= 10 \times \epsilon^3 \times (1-\epsilon)^2 + 5 \times \epsilon^4 \times (1-\epsilon) + \epsilon^5
\end{aligned}
$$

### (2)

$$
\begin{aligned}
E &= 10 \times (0.2)^3 \times (1-0.2)^2 + 5 \times (0.2)^4 \times (1-0.2) + (0.2)^5 \\
&= 10 \times 0.008 \times 0.64 + 5 \times 0.0016 \times 0.8 + 0.00032 \\
&= 0.0512 + 0.0064 + 0.00032 \\
&= 0.05792
\end{aligned}
$$

# Task 4

**Technical report in another pdf file, please check Technical_Report.pdf**