

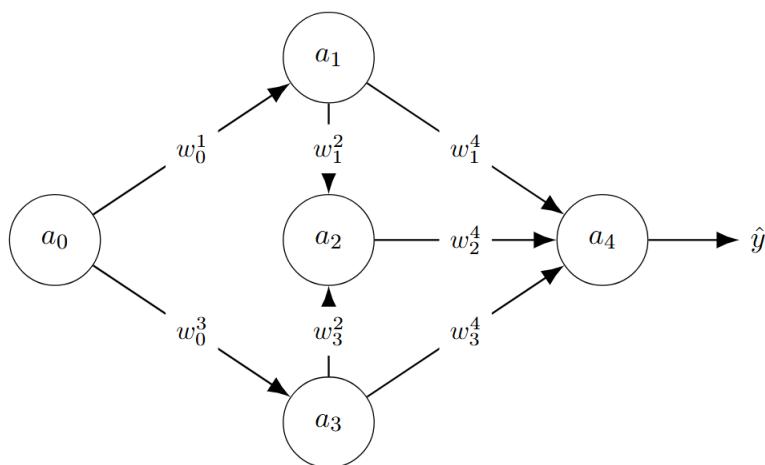
**BRAC UNIVERSITY**  
**CSE422 : Artificial Intelligence**  
**Assignment 4**

---

1. Consider the following dataset, where  $x_1$  and  $x_3$  given  $y$  belong to a multinomial distribution and  $x_2$  given  $y$  belongs to a normal distribution:

$x_1$	$x_2$	$x_3$	$y$
0	28.5	1	0
0	29.2	1	0
1	25.4	1	1
2	21.1	1	1
2	18.5	0	1
2	19.2	0	0
1	22.8	0	1
0	27.6	1	0
0	28.9	0	1
2	21.8	0	1

- a. Calculate the model parameters for Naive Bayes.
  - b. Calculate  $p(y = 1 | x_1 = 1, x_2 = 25.2, x_3 = 0)$ .
  - c. Infer the value of  $y$  given  $x_1 = 2, x_2 = 22.2, x_3 = 1$ .
  - d. Calculate the model parameters using Laplace smoothing.
2. Show that a simple perceptron cannot solve the XOR problem. The general equation of a perceptron with two inputs is given by:  $w_0 + w_1X_1 + w_2X_2 = 0$ . Assume that there exists a solution and show that this leads to a contradiction.
3. Consider the following neural network:



Here  $a_i = f_i(\sum_j w_j^{[i]} a_{i-1})$  for  $i \in \{1, 2, 3, 4\}$ ,  $a_0 = x$  (input neuron),  $f_2(x) = \text{ReLU}(x)$ , and  $f_1(x) = f_3(x) = f_4(x) = \sigma(x)$ . Let the square loss be the loss function and the initial value of the parameters  $w_0^1 = 1, w_0^3 = 2, w_1^2 = 1.5, w_3^2 = 1, w_1^4 = 2.5, w_2^4 = 1.2, w_3^4 = 0.8$ .

- a. Write a function that takes the input  $x$  and computes the neural network output.
- b. Deduce the equations to calculate the derivative of the loss function with respect to  $w_3^4, w_3^2, w_0^1$ .
- c. Calculate the output of the neural network for  $x = 0.7$  and  $1.5$ .
- d. Update the parameters  $w_3^4, w_3^2, w_0^1$  for 2 epochs and learning rate 0.1 with the following training set:  $\{(0, 0.98), (-0.5, 0.89), (1.2, 0.52)\}$
4. You are to build a decision tree to predict whom, among candidate A, B, and C, a person will vote in an election based on their age group (18-30, 31-50, 51+), income level (Low, Medium, High), and interest in politics (Low, Medium, High). The dataset contains the following distribution:
- 50 people in the 18-30 group: 25 vote for A, 15 vote for B, and 10 vote for C.
  - 40 people in the 31-50 group: 30 vote for A, 5 vote for B, and 5 vote for C.
  - 30 people in the 51+ group: 5 votes for A, 10 vote for B, and 15 vote for C.
- a. Compute the Entropy of the given dataset.
- b. Compute the Conditional entropy of Age Group.
- c. Compute the information gain if you were to split the data based on the age group.
- Consider the following:
- 60 people have low interest: 10 vote for A, 40 vote for B, and 10 vote for C.
  - 30 people have medium interest: 20 vote for A, 5 vote for B, and 5 vote for C.
  - 30 people have high interest: 30 vote for A, 0 vote for B, and 0 vote for C.
- d. Calculate the information gain if you split the dataset based on "Interest in Politics."
- e. Between "Interest in Politics" and "Age Group" which attribute should be in the root node of the decision tree.