FIT5215 Deep Learning

Quiz for:
Feed-forward Neural Nets with PyTorch

Trung Le and **Teaching Team**

Department of Data Science and AI
Faculty of Information Technology, Monash University
Email: **trunglm@monash.edu**

Question 1

- *Assume that we have **4 classes** in **{cat = 1,dog = 2,lion = 3, monkey = 4}**. Given a data example x with **ground-truth label "dog"**, assume that a feed-forward NN gives **discriminative scores** to this x as $h_1 = 7$, $h_2 = 10$, $h_3 = 5$, $h_4 = -2$. Choose all correct answers. (MC)*

- A. The model predicts x as cat
- B. The model predicts x as dog
- C. This is a correct prediction
- D. This is an incorrect prediction

Question 2

- *Assume that we have 4 classes in {cat = 1,dog = 2,lion = 3, monkey = 4}. What is one-hot label of categorical label "**dog**"?*

- A. [1,0,0,0]
- B. [0,1,0,0]
- C. [0,0,1,0]
- D. [0,0,0,1]

Question 3

- *Assume that we have **4 classes** in **{cat = 1,dog = 2,lion = 3, monkey = 4}**. Given a data example x with **ground-truth label "dog"**, assume that a FFN model gives **discriminative scores** to this x as $h_1 = -3$, $h_2 = 10$, $h_3 = 5$, $h_4 = 0$. What is the probability to predict x as dog or $p(y = dog \mid x)$?*

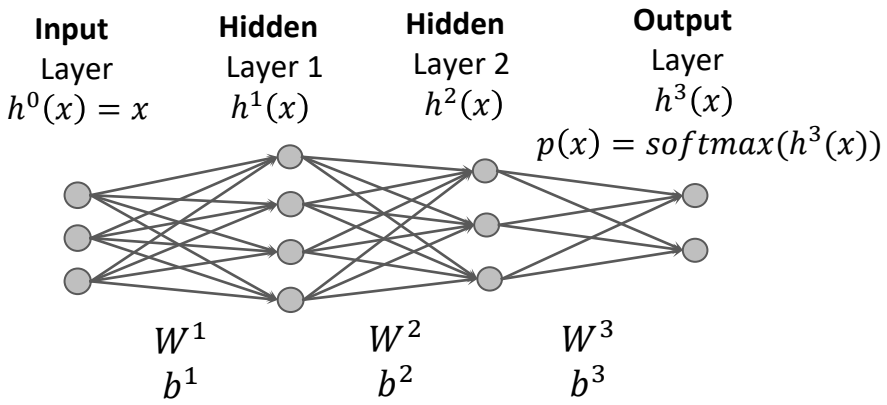
- A. $\frac{e^5}{e^{-3}+e^{10}+e^5+e^0}$
- B. 1
- C. $\frac{e^0}{e^{-3}+e^{10}+e^5+e^0}$
- D. $\frac{e^{10}}{e^{-3}+e^{10}+e^5+e^0}$

Question 4

- *Assume that we have **4 classes** in **{cat = 1,dog = 2,lion = 3, monkey = 4}**. Given a data example x with **ground-truth label "dog"**, assume that a feed-forward NN gives **discriminative scores** to this x as $h_1 = -3$, $h_2 = 10$, $h_3 = 5$, $h_4 = 0$. What is the CE loss suffered by this prediction?*
- A. $-\log \frac{-3}{e^{-3}+e^{10}+e^{5}+e^{0}}$
- B. $\log \frac{e^{h_3}}{\sum_{j=1}^{4} e^{h_j}}$
- C. $-\log \frac{e^{10}}{e^{-3}+e^{10}+e^{5}+e^{0}}$
- D. $\log \frac{e^{-h_3}}{\sum_{j=1}^{4} e^{h_j}}$

Question 5

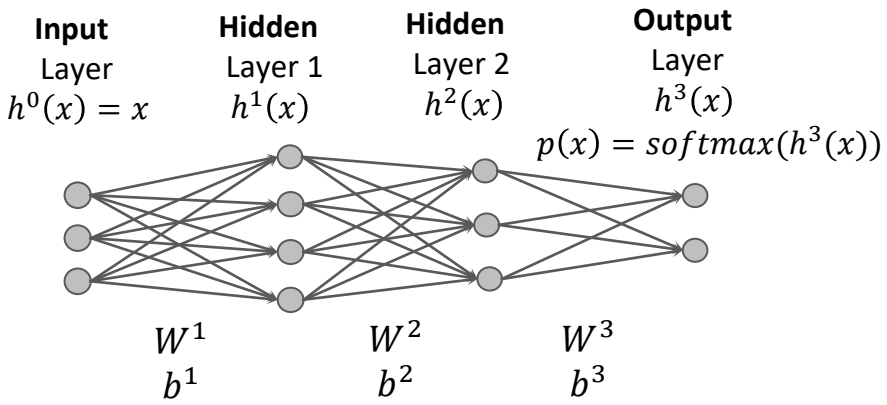
- *Given the following feed-forward neural network. What are the shapes of weight matrices?*



- A. $W^1 \in \mathbb{R}^{3\times4}, W^2 \in \mathbb{R}^{4\times3}, W^3 \in \mathbb{R}^{3\times2}$
- B. $W^1 \in \mathbb{R}^{4\times3}, W^2 \in \mathbb{R}^{3\times4}, W^3 \in \mathbb{R}^{2\times3}$
- C. $W^1 \in \mathbb{R}^{3\times4}, W^2 \in \mathbb{R}^{4\times3}, W^3 \in \mathbb{R}^{3\times2}$
- D. $W^1 \in \mathbb{R}^{4\times4}, W^2 \in \mathbb{R}^{3\times3}, W^3 \in \mathbb{R}^{2\times2}$

Question 6

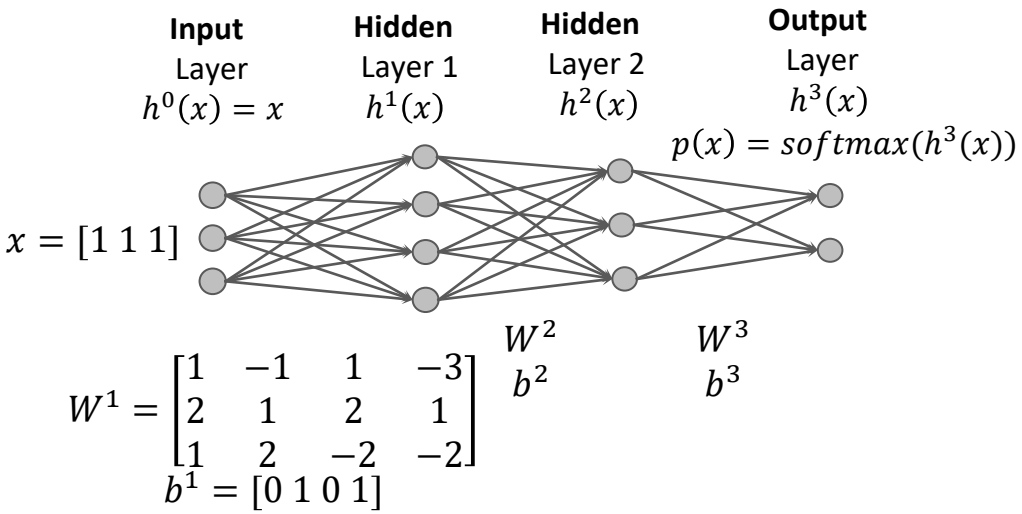
- *Given the following feed-forward neural network. We feed the mini-batch $x \in \mathbb{R}^{16\times 3}$ to the network. What is the shape of the hidden values h^2?*



- A. [16, 4]
- B. [16, 3]
- C. [3, 16]
- D. [4, 16]

Question 7

- *Given the following feed-forward neural network. Assume that we input to the network feature vector* $x = [1\ 1\ 1]$. *What is the values of pre-activations* $\bar{h}^1$ *?*

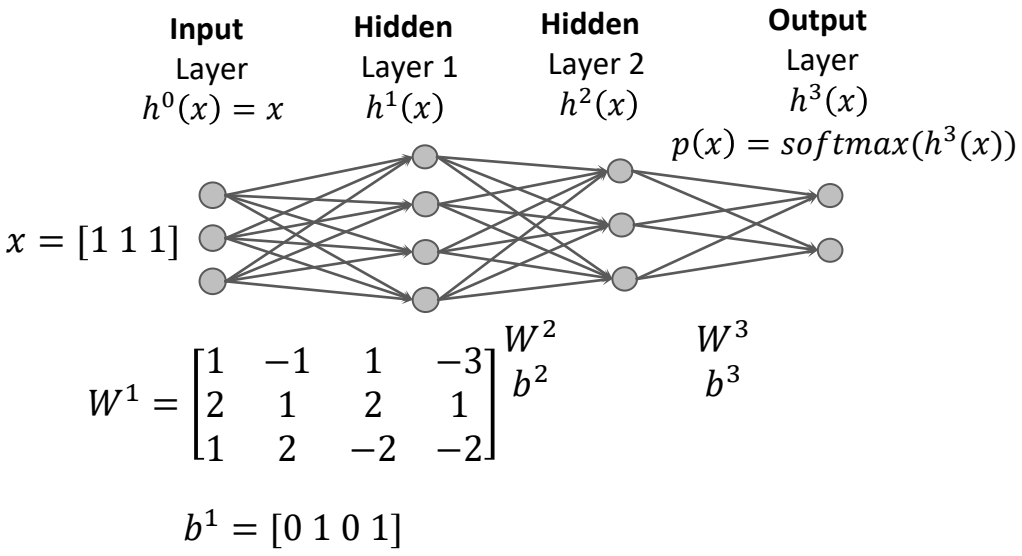


- A. $\bar{h}^1 = [4\ 2\ 1\ -4]$
- B. $\bar{h}^1 = [4\ 3\ 1\ -3]$
- C. $\bar{h}^1 = [4\ 3\ 1\ -3]^T$
- D. $\bar{h}^1 = [4\ 2\ 1\ -4]^T$

Question 8

- *Given the following feed-forward neural network. Assume that we input to the network feature vector* $x = [1\ 1\ 1]$. *What is the values of hidden values* h^1 *if we use activation ReLU?*

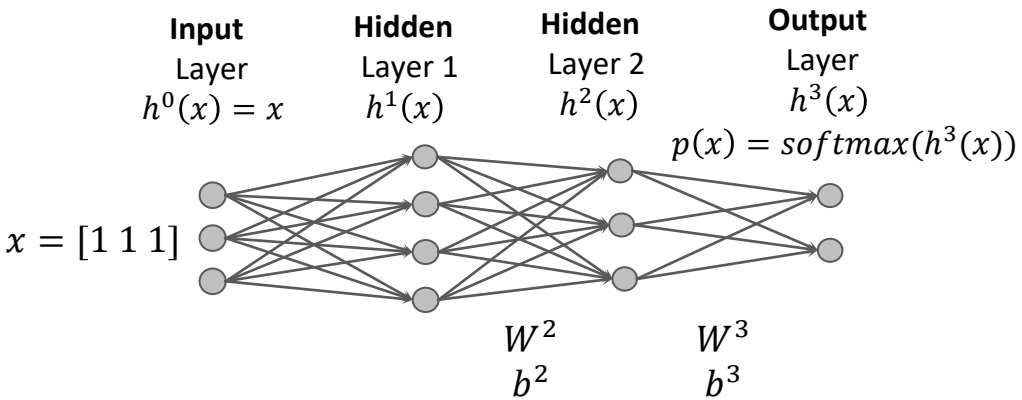
- A. $h^1 = [4\ 2\ 1\ 0]$
- B. $h^1 = [4\ 3\ 1\ 0]$
- C. $h^1 = [4\ 3\ 1\ 0]^T$
- D. $h^1 = [4\ 2\ 1\ 0]^T$



Question 9

- *Given the following feed-forward neural network. Assume that we input to the network feature vector* $x = [1\ 1\ 1]$. *What is the values of hidden values* h^1 *if we use activation ReLU?*

- A. $\mathrm{h}^1 = [1\ 1\ 1\ 0]$
- B. $\mathrm{h}^1 = [1\ 1\ 1\ -2]$
- C. $\mathrm{h}^2 = [0\ \ 2\ \ -1]$
- D. $\mathrm{h}^2 = [0\ \ 2\ \ 0]$



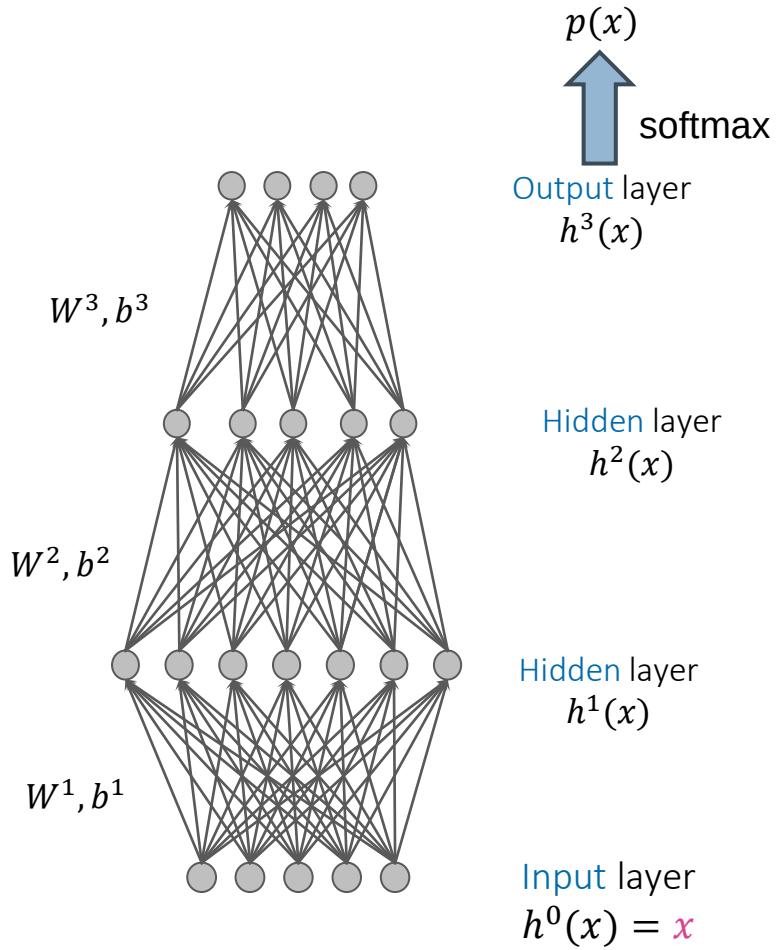
$$W^1 = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 0 & 0 & -1 \\ 1 & 1 & 2 & 1 \end{bmatrix} \quad W^2 = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 2 & -2 & 2 \end{bmatrix}$$

$$b^1 = [0\ 1\ 0\ -1] \qquad b^2 = [-1\ 1\ -2]$$

Question 10

- *Given the following feed-forward neural network. Assume that we input to the network a mini-batch x with the batch size 32 What is the shape of the input x if we use activation ReLU?*

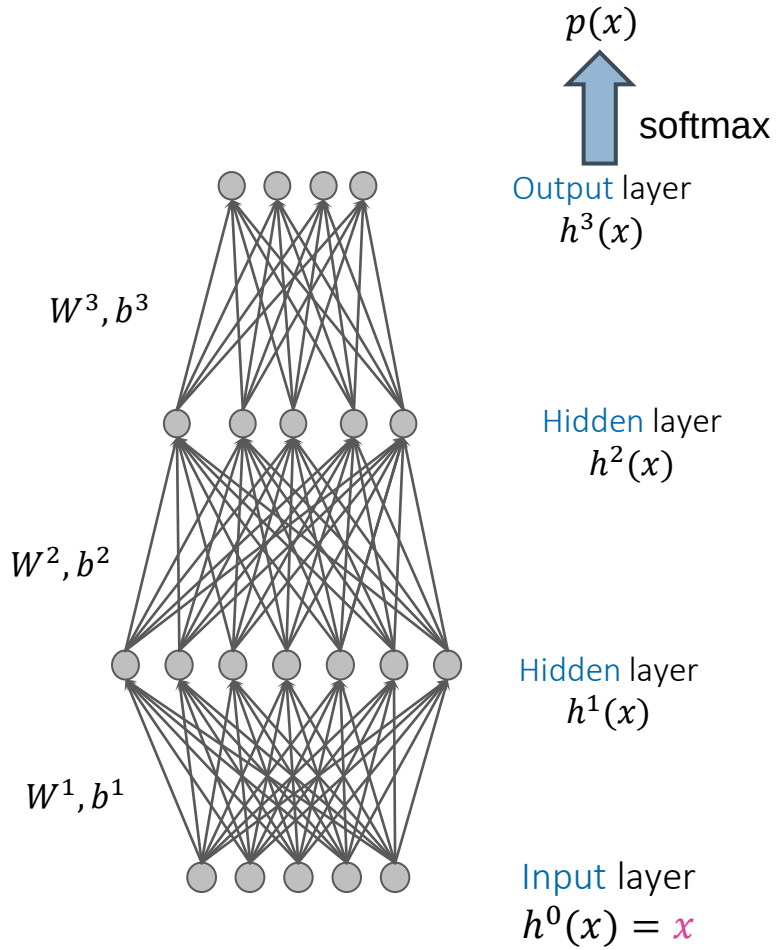
- A. [32,5]
- B. [5,32]
- C. [32,4]
- D. [5,4]



Question 11

- *Given the following feed-forward neural network. Assume that we input to the network mini-batch $x \in \mathbb{R}^{32\times5}$. What is the shape of the logits h^3 if we use activation ReLU?*

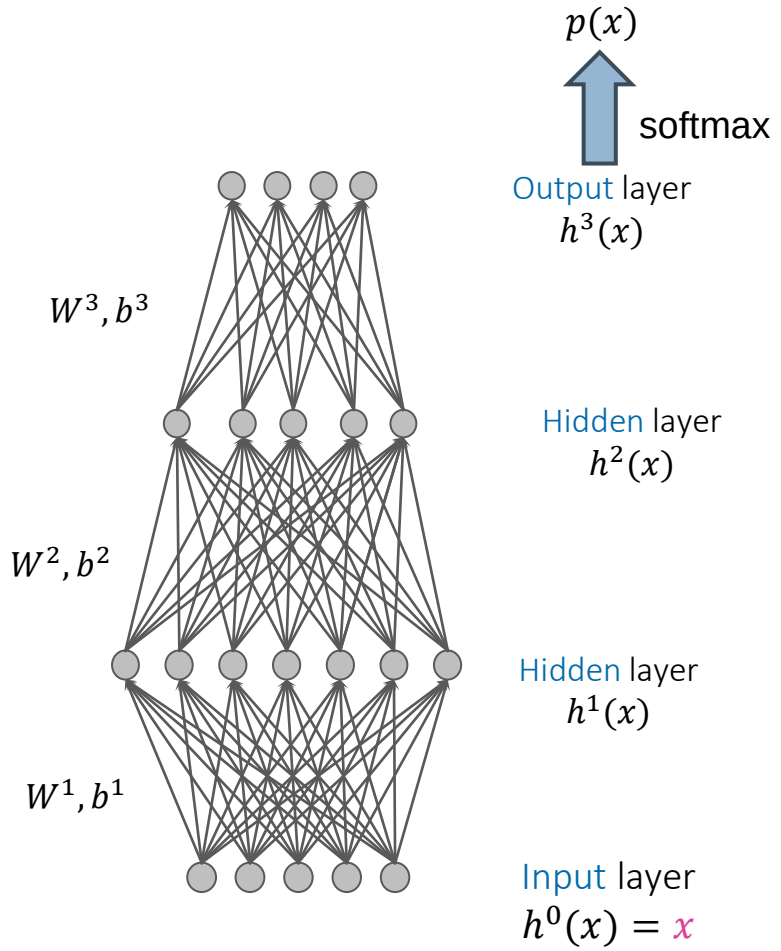
- A. [32,5]
- B. [5,32]
- C. [32,4]
- D. [5,4]



Question 12

- *Given the following feed-forward neural network. Assume that we input to the network mini-batch $x \in \mathbb{R}^{32\times5}$. What is the meaning of the 5th row in the logits h^3?*

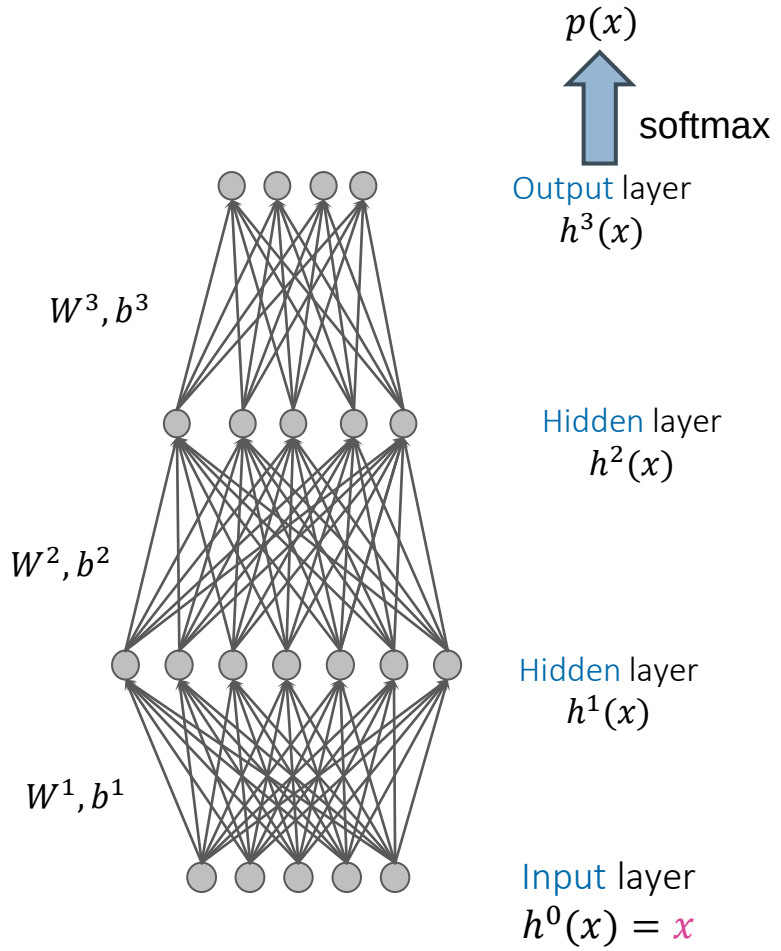
- A. It has no meaning
- B. It is the 5th logit values of all data points in the mini-batch.
- C. It is the logits of the 5th data point in our batch.
- D. None of above.



Question 13

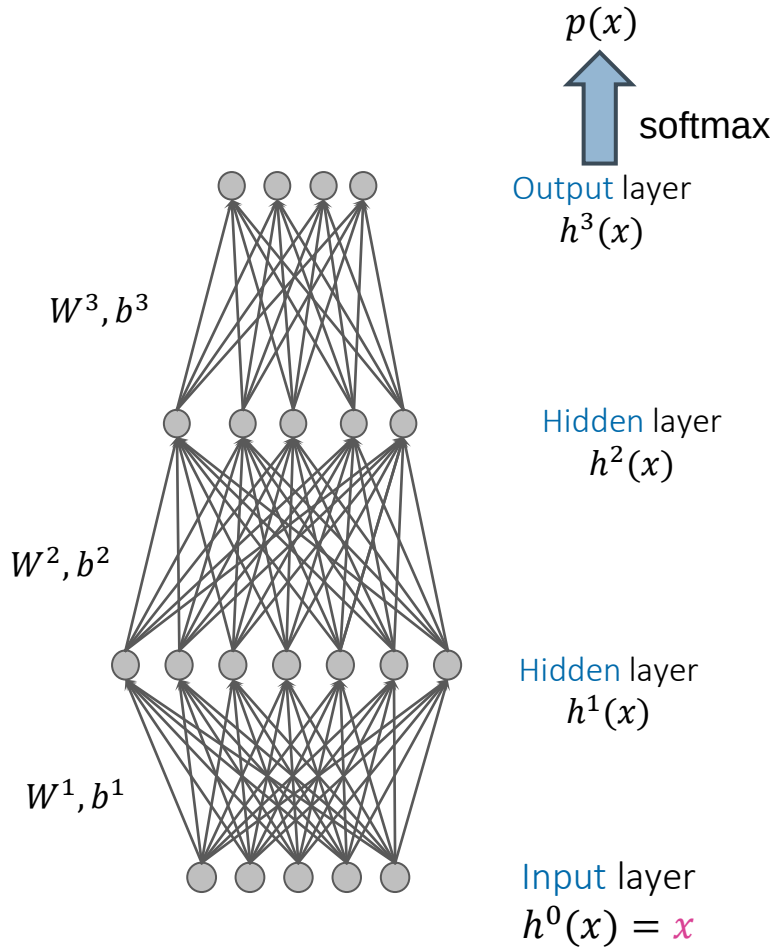
- *Given the following feed-forward neural network. Assume that we input to the network mini-batch $x \in \mathbb{R}^{32\times5}$. What is the shape of the prediction probabilities p if we use activation ReLU?*

- A. [32,5]
- B. [5,32]
- C. [32,4]
- D. [5,4]



Question 14

- *Given the following feed-forward neural network. Assume that we input to the network mini-batch $x \in \mathbb{R}^{32\times5}$. What is the meaning of the 4th row in the prediction probabilities p?*



- A. It has no meaning
- B. It is the 4th probabilities values of all data points in the mini-batch.
- C. It is the prediction probabilities of the 4th data point in our batch.
- D. None of above.

Question 15

- *Given an implementation as below (assume that **n_features =16**). Which of following architecture is correct (SC).*

```
device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")
dnn_model = Sequential(Linear(n_features,10), nn.ReLU(),
                       Linear(10,20), nn.ReLU(),
                       Linear(20,15), nn.ReLU(),
                       Linear(15, 26)).to(device)
```

- A. 16→10(ReLU)→20(ReLU)→15(ReLU)→26(ReLU)
- B. 16→10(ReLU)→20(ReLU)→15(ReLU)→26(ReLU) (Softmax)
- C. 16→10(ReLU)→20(ReLU)→15(ReLU)→26
- D. 16→10(sigmoid)→20(sigmoid)→15(sigmoid)→26(sigmoid)

Question 16

- *Given an implementation as below (assume that **n_features =16**). What is the total number parameters of this FFN (SC)?*

```
device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")
dnn_model = Sequential(Linear(n_features,10), nn.ReLU(),
                       Linear(10,20), nn.ReLU(),
                       Linear(20,15), nn.ReLU(),
                       Linear(15, 26)).to(device)
```

- A. 16x10 + 10x20 + 20x15 + 15x26
- B. 16x10 +10 + 10x20 + 20 + 20x15 +15 + 15x26 +26
- C. 16x10 +16 + 10x20 + 10 + 20x15 +20 + 15x26 +15
- D. None of above.

Question 17

- *Given the code as below (assume that **n_features =16** and **n_classes=26**). What is the shape of the logits h (SC)?*

```
device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")
dnn_model = Sequential(Linear(n_features,10), nn.ReLU(),
                       Linear(10,20), nn.ReLU(),
                       Linear(20,15), nn.ReLU(),
                       Linear(15, n_classes)).to(device)

x = torch.rand(32,n_features).to(device)
h = dnn_model(x)
print(h.shape)
```

- A. [32,16]
- B. [32,26]
- C. [32,15]
- D. None of above.