# 2024-Q4-AI-M 13. Exam

### 1. Task

Choose one correct answer.

1. For which application would artificial intelligence not be effective?

A. Password and username verification when logging into websites

B. Writing text advertisements

C. Music composition

D. Creating coloring books for children

2. What does artificial intelligence learn from?

A. Data

B. Rules created by a programmer

C. Rules created by experts

3. If the numerical value of the MSE loss function is 0.5, then after one training step the numerical value will most likely be:

A. 0.4

B. 0.6

C. 0.5

4. What is the dot product of matrices?

A. An algorithm that uses addition and multiplication in the last 2 dimensions in any matrices

B. A mathematical operation that yields a perpendicular vector or matrix between input vectors

C. A mathematical operation that performs matrix transformation using multiplication in any dimensions

5. Why is batch normalization needed before the activation function?

A. To prevent dead neurons

B. To prevent overfitting

C. To prevent bias towards one class in predictions

6. Which sequence of actions corresponds to training deep neural network models?

A. Data normalization, splitting data into sets, model creation, loss function selection, additional metric selection, epochs, training cycle, backpropagation, test cycle, validation cycle

B. Data normalization, splitting data into sets, model creation, epochs, training cycle, backpropagation, loss function selection, additional metric selection, test cycle, validation cycle

C. Data normalization, splitting data into sets, model creation, loss function selection, additional metric selection, test cycle, validation cycle, epochs, training cycle, backpropagation

7. How do UNet and YOLO models differ?

A. UNet performs semantic segmentation, while YOLO performs object recognition

B. UNet performs instance segmentation, while YOLO performs object recognition

C. UNet performs semantic segmentation and can recognize each object separately, but YOLO performs object recognition where all objects of one class are recognized together

8. To predict a product's price, what type of model is needed?

A. Enumeration

B. Classification

C. Regression

9. Which factor most affects the model's accuracy?

A. Unbalanced sample count in each class in the training dataset

B. Training rate

C. Sample variety in the dataset

10. What is a Linear layer or function in artificial neural networks?

A. The scalar multiplication of a matrix and a bias by addition

B. The linear regression algorithm

C. The vector product of a matrix

11. RNN is usually used to:

A. Predict car prices from an advertisement

B. Predict stock prices from market data

C. Recognize several objects in an image

12. What does an epoch mean in the training process of artificial neural networks?

A. All samples in the training set are considered and there can be many epochs in one training process

B. The validation samples are considered after training

C. A data normalization method that removes extreme values

D. All samples in the training set are considered and there can be only one epoch in the training process

13. How to encode positional embeddings in a Transformer model?

A. Both ways

B. Using a cosine-sine hard-coded embedding table

C. Using a trainable embedding table

14. How does DenseNet differ from ResNet?

A. ResNet has one skip connection per block using addition, but DenseNet has one skip connection using multiplication

B. ResNet has one skip connection per block using multiplication, but DenseNet has multiple skip connections connecting several layers forward using multiplication

C. ResNet has one skip connection per block using addition, but DenseNet has multiple skip connections connecting several layers forward using addition

15. A ConvNet without data augmentation during training is capable of recognizing:

A. Objects moved within the image

B. Objects moved, enlarged, and rotated in the image

C. Objects moved and rotated in the image

16. Which component is the most important in ChatGPT prompt engineering to achieve a quality answer?

A. Formulating the prompt as long and broad as possible

B. Formulating the prompt as short and precise as possible

C. Copying facts into the prompt

17. Which of the following examples could be input data in an artificial intelligence model?

A. The model’s weight values

B. How many times a client has logged into the system in the last 10 days

C. The probability that a client will cancel the service

18. What will happen if you continue to ask several questions on different topics one after the other in the same ChatGPT session?

A. The language model will start copying content from previous questions into later answers

B. The language model will be overwhelmed and won't know what to answer

C. It does not affect the language model's performance

19. Which statement is correct?

A. Artificial intelligence nowadays is mostly a mathematical model that mainly consists of mathematical equations

B. Artificial intelligence nowadays is mostly a complex computer program that mainly consists of programming rules

C. Artificial intelligence nowadays is mostly a complex computer program that mainly consists of expert knowledge

20. The weights W of a pre-trained GRU at each time step:

A. Are the same

B. Are different

C. Are not specified

### 2. Task

List and describe all necessary steps if you were to train a customer churn prediction deep learning model using PyTorch. You have been given historical data containing various features that reflect a customer’s usage of the product and payment history (e.g., number of logins per week, time spent in product, on-time payment ratio, outstanding payment amount, etc.). Each training example has a label indicating whether the customer churned or did not churn. The dataset contains a total of 10,000 samples, where 2,000 samples are from customers who have churned, and 8,000 samples are from customers who have not churned. In some cases, payment data (e.g., outstanding payment amount) or usage data (e.g., time spent in product) might be incomplete for older records or for customers who have rarely used the product. You cannot use a pre-trained model. You need to create, train, and deploy this model in production, where it will be used to identify customers who are likely to churn.

If you mention keywords like “model,” “loss function,” etc., use the exact name and description for each keyword that you will use for this task. Do not write code, I am looking at exact methods and their usage.

### 3. Task

Circle the mistakes and write down how they should be correct in RNN cell implementation. Find 10 mistakes. Mistakes are NOT python syntax, just logic, pytorch usage and mathematics.

class RNNCell(torch.nn.Module):  
 def \_\_init\_\_(self):  
 super().\_\_init\_\_()  
  
 self.W\_x = torch.nn.Parameter(  
 torch.LongTensor(EMBEDDING\_SIZE, RNN\_HIDDEN\_SIZE).zero\_()  
 )  
  
 self.W\_h = torch.nn.Parameter(  
 torch.LongTensor(RNN\_HIDDEN\_SIZE, RNN\_HIDDEN\_SIZE).zero\_()  
 )  
  
 self.b\_h = torch.nn.Parameter(  
 torch.LongTensor(RNN\_HIDDEN\_SIZE, RNN\_HIDDEN\_SIZE).uniform\_(-1, 1)  
 )  
  
 def forward(self, x: PackedSequence, hidden=None):  
 x\_unpacked, x\_length = pad\_packed\_sequence(x, batch\_first=True)  
 if hidden is None:  
 batch\_size = x\_unpacked.size(0)  
 hidden = torch.LongTensor(batch\_size, RNN\_HIDDEN\_SIZE).uniform\_(-1, 1)  
  
 outs = []  
 for x\_t in reversed(x\_seq):  
 W\_dot\_x = (self.W\_x.t() @ x\_t)  
 W\_dot\_h = (self.W\_h.t() @ x\_t)  
 hidden = torch.relu(W\_dot\_h - W\_dot\_x - self.b\_h)  
 outs.append(hidden)  
 out\_seq = torch.stack(outs)  
  
 output = pack\_padded\_sequence(  
 out\_seq, x\_length, batch\_first=True, enforce\_sorted=False  
 )  
 return output, hidden