# 2024-Q4-AI-Business 9. Exam

Circle *one* correct answer!

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

A. Copying facts into the prompt

B. Formulating the prompt as short and precise as possible

C. Formulating the prompt as long and vague as possible

2. For what purposes could clustering and categorization using the k-Means algorithm be useful?

A. Generating text for advertisements

B. Recommending new products to customers using their purchase history

C. Predicting the price of a new product that differs significantly from all existing products

D. Classifying animal images

3. What do Large Language Models (LLM) resemble the most?

A. An improvisational theater that responds based on the information provided by the user

B. An Internet search engine

C. An oracle that can answer all questions

4. What is the purpose of using metrics such as F1 or Accuracy when training an artificial intelligence model?

A. F1 and Accuracy serve as the main criteria to compare different AI models with each other and choose the best model for a given task. The higher these metrics, the better the model.

B. F1 and Accuracy are necessary to adjust the AI model's hyperparameters and improve its architecture. Based on these metrics, one can understand how to change the model's structure.

C. F1 and Accuracy are used to check the speed and efficiency of the AI model during training. The faster the model trains, the better these metrics are.

D. F1 and Accuracy are used to evaluate the AI model's performance and quality during training and testing. They help understand how well the model can classify or predict outcomes correctly.

5. In what format do you encode categorical data in artificial intelligence to train a model, for example: bmw, audi, toyota?

A. As class indices: 0 = bmw, 1 = audi, 2 = toyota

B. One-hot-encoded: [1, 0, 0] = bmw, [0, 1, 0] = audi, [0, 0, 1] = toyota

C. Both ways

6. To use a Large Language Model (LLM) most effectively with company data, what is needed:

A. Train on company data

B. Connect a text semantic meaning model and create a RAG (Retrieval Augmented Generation) system that uses a pretrained model

C. Program the model to search for company data in the database by itself

7. What is the purpose of using a loss function when training artificial intelligence?

A. The loss function is used to evaluate the difference between the AI's predicted and actual values, allowing the model to adjust for better accuracy.

B. The loss function is used to check the operation of AI hardware and identify possible errors.

C. The loss function is used to measure the speed and efficiency of AI operations, determining its performance.

D. The loss function serves as a safety mechanism to prevent the AI from having too much autonomy and independent operation.

8. Where is PCA (Principal Component Analysis) useful?

A. None of the given options

B. PCA is used to increase the dimensionality of input data to visualize data and see their relationships

C. PCA is useful for reducing the dimensionality of input data, in order to visualize data and see their relationships

9. The Transformer model, which underlies ChatGPT, is based on:

A. Programming and statistical rules to find necessary text in database

B. An attention mechanism that pays attention to the input text

C. On a text database

10. How does Apple FaceID work, which uses a face re-identification model?

A. The model is already pretrained and obtains a unique embedding vector

B. Both options

C. The model is trained using the user's face data

11. What are histograms?

A. A histogram is a mathematical formula used to calculate the average value of a dataset, taking into account the deviation of the data from the average.

B. A histogram is a statistical method used to test whether the difference between two datasets is statistically significant, comparing their means and standard deviations.

C. A histogram is a graphical representation that shows the distribution of data by dividing them into several intervals or bins and representing the frequency of each interval as the height of a bar.

D. A histogram is a type of data visualization in which data is represented as a line graph, where each data point is connected by a straight line to show changes over time.

12. Which factors have contributed most to the development of artificial intelligence in the last 10 years?

A. Data availability, computing resource power, mathematical theory

B. Business applications, data availability, computing resource power

C. Public interest, business applications, computing resource power

13. In what format should the input data be to train an AI model to predict apartment prices from advertisement text?

A. Both ways

B. Absolute values: 0…500k EUR

C. Normalized price ranges: -1..1

14. What type of model is required to predict the price of a product?

A. Enumeration

B. Classification

C. Regression

15. What type of input data is used in STT or ASR models?

A. Audio signals

B. Video data

C. Image data

D. Text data

16. Which application would artificial intelligence not be effective for?

A. Composing music

B. Checking passwords and usernames during website authentication

C. Writing text advertisements

D. Creating coloring books for children

17. Where is voice recording enhancement (speech enhancement) with artificial intelligence useful?

A. In speech recognition and transcription

B. In cases when a high-quality voice recording is available

C. In forensic analysis and criminology

18. What does an Epoch mean in the training process of artificial neural networks?

A. All samples in the training set are reviewed, and there can only be one epoch in the training process

B. A data normalization method that removes extreme values

C. All samples in the training set are reviewed, and there can be many epochs in one training process

D. Validation samples are reviewed after training

19. What will happen if you continue asking several questions on different topics consecutively in the same ChatGPT session?

A. It will not affect the language model's performance

B. The language model will become confused and not know what to answer

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

20. Which of the given examples could be output data in an artificial intelligence model?

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

B. The values of the model's weights

C. The probability that a client will cancel a service