Claims

1. A data processing method for a task processing model, applied to a cloud-side device, wherein the cloud-side device is connected with a plurality of terminal-side devices, and the method comprises:

   obtaining a first sample set, wherein the first sample set comprises multi-modal sample guiding information;

   inputting the sample guiding information and a sample task sequence into an initial processing model to obtain predicted task embeddings corresponding to the sample guiding information;

   training the initial processing model according to the predicted task embeddings and sample task embeddings corresponding to the sample task sequence to obtain model parameters of a task processing model in response to reaching a first preset stopping condition, wherein the sample task embeddings are obtained by quantizing and encoding the sample task sequence; and

   sending the model parameters of the task processing model to the plurality of terminal-side devices.

2. The method according to claim 1, wherein the initial processing model comprises a first encoder and a first decoder, and the first preset stopping condition comprises a first stopping sub-condition, inputting the sample guiding information and the sample task sequence into the initial processing model to obtain the predicted task embeddings corresponding to the sample guiding information comprises:

   extracting first sample guiding information from the first sample set, wherein the first sample guiding information is any sample guiding information in the first sample set;

   inputting the first sample guiding information into the first encoder to obtain first sample guiding embeddings;

   masking the sample task sequence to obtain a masked task sequence; and

   inputting the first sample guiding embeddings and the masked task sequence into the first decoder to obtain first predicted task embeddings; and

   training the initial processing model according to the predicted task embeddings and the sample task embeddings corresponding to the sample task sequence to obtain the model

parameters of the task processing model in response to reaching the first preset stopping condition comprises:

calculating a decoding loss value based on the first predicted task embeddings and the sample task embeddings corresponding to the sample task sequence; and

adjusting parameters of the first decoder according to the decoding loss value, and returning to execute a step of extracting the first sample guiding information from the first sample set to obtain the first decoder in response to reaching the first stopping sub-condition.

3. The method according to claim 2, wherein the first preset stopping condition comprises a second stopping sub-condition, and a training method for the first encoder comprises:

obtaining a second sample set, wherein the second sample set comprises multi-modal sample guiding information carrying sample guiding embeddings;

extracting second sample guiding information from the second sample set, wherein the second sample guiding information is any sample guiding information in the second sample set;

inputting the second sample guiding information into the first encoder to obtain first predicted guiding embeddings corresponding to the second sample guiding information;

calculating an encoding loss value based on the first predicted guiding embeddings and second sample guiding embeddings carried in the second sample guiding information; and

adjusting parameters of the first encoder according to the encoding loss value, and returning to execute a step of extracting the second sample guiding information from the second sample set to obtain the first encoder in response to reaching the second stopping sub-condition.

4. The method according to claim 2, before calculating the decoding loss value based on the first predicted task embeddings and the sample task embeddings corresponding to the sample task sequence, further comprising:

inputting the sample task sequence into a second encoder of a pre-trained quantized generation model to obtain the sample task embeddings corresponding to the sample task sequence, after encoding the sample task sequence by the second encoder.

5. The method according to claim 2, wherein inputting the first sample guiding embeddings and the masked task sequence into the first decoder to obtain the first predicted task embeddings comprises:

obtaining a preset timestep embedding, wherein the timestep embedding corresponds one-to-one with each of training iterations; and

inputting the timestep embedding, the first sample guiding embeddings and the masked task sequence into the first decoder to obtain the first predicted task embeddings.

6. The method according to claim 4, wherein the quantized generation model comprises a second encoder and a second decoder, and the training method for the quantized generation model comprises:

obtaining a third sample set, wherein the third sample set comprises a plurality of training sample sequences;

extracting a first training sample sequence from the third sample set, wherein the first training sample sequence is any training sample sequence in the third sample set;

inputting the first training sample sequence into the second encoder to obtain first test embeddings;

inputting the first test embeddings and preset prompt information into the second decoder to obtain a first test sequence;

calculating a quantization loss value based on the first training sample sequence and the first test sequence;

adjusting parameters of the second encoder and parameters of the second decoder according to the quantization loss value, and returning to execute a step of extracting the first training sample sequence from the third sample set to obtain the model parameters of the trained quantized generation model in response to reaching the second preset stopping condition; and

sending the model parameters of the trained quantized generation model to the plurality of terminal-side devices.

7. The method according to claim 6, after inputting the first test embeddings and the preset prompt information into the second decoder to obtain the first test sequence, further comprising:

randomly slicing the first test sequence to obtain a first test subsequence and a second test subsequence;

calculating a continuity loss value based on the first test subsequence and the second test subsequence;

adjusting the parameters of the second encoder and the parameters of the second decoder according to the quantization loss value, and returning to execute the step of extracting the first training sample sequence from the third sample set to obtain the model parameters of the trained quantized generation model in response to reaching the second preset stopping condition comprises:

adjusting the parameters of the second encoder and the parameters of the second decoder according to the quantization loss value and the continuity loss value, and returning to execute the step of extracting the first training sample sequence from the third sample set to obtain the model parameters of the trained quantized generation model in response to reaching the second preset stopping condition.

8. A data processing method for a task processing model, applied to a terminal-side device, wherein the terminal-side device is connected with a cloud-side device and the method comprises:

receiving model parameters of a task processing model sent by the cloud-side device, and constructing a task processing model according to the model parameters;

receiving a task processing request input by a user, wherein the task processing request comprises task guiding information;

inputting the task guiding information and a fully-masked task sequence into the task processing model to obtain target task embeddings corresponding to the task guiding information after processing the task guiding information and the fully-masked task sequence by the task processing model, wherein the task processing model is trained based on multi-modal sample guiding information, a sample task sequence and sample task embeddings, and the sample task embeddings are obtained by quantizing and encoding the sample task sequence; and

quantizing and decoding the target task embeddings to obtain a task processing result corresponding to the task guiding information.

9. The method according to claim 8, wherein the task processing model comprises a first encoder and a first decoder, inputting the task guiding information and the fully-masked task sequence into the task processing model to obtain the target task embeddings corresponding to the task guiding information after processing the task guiding information and the fully-masked task sequence by the task processing model comprises:

inputting the task guiding information into the first encoder to obtain task guiding embeddings corresponding to the task guiding information; and

inputting the task guiding embeddings and the fully-masked task sequence into the first decoder to obtain the target task embeddings corresponding to the task guiding information.

10. A method for generating virtual character animation, comprising:

receiving a virtual character animation generation request sent by a frontend, wherein the virtual character animation generation request comprises animation guiding information;

inputting the animation guiding information and a fully-masked animation sequence into a virtual character animation generation model to obtain virtual character animation embeddings corresponding to the animation guiding information after processing the animation guiding information and the fully-masked animation sequence by the virtual character animation generation model, wherein the virtual character animation generation model is trained based on sample animation guiding information, a sample animation sequence, and sample animation embeddings, and the sample animation embeddings are obtained by quantizing and encoding the sample animation sequence;

quantizing and decoding the virtual character animation embeddings to obtain a virtual character action sequence corresponding to the animation guiding information; and

generating virtual character animation based on the virtual character action sequence and sending the virtual character animation to the frontend to display the virtual character animation.

11. The method according to claim 10, before inputting the animation guiding information and the fully-masked animation sequence into the virtual character animation generation model to obtain the virtual character animation embeddings corresponding to the animation guiding information after processing the animation guiding information and the fully-masked animation sequence by the virtual character animation generation model, further comprising:

receiving designated prompt information input by a user, wherein the designated prompt information comprises at least one of a designated prompt text and a designated prompt audio; and

a step of inputting the animation guiding information and the fully-masked animation sequence into the virtual character animation generation model to obtain the virtual character animation embeddings corresponding to the animation guiding information after processing the animation guiding information and the fully-masked animation sequence by the virtual character animation generation model comprises:

inputting the designated prompt information, the animation guiding information and the fully-masked animation sequence into the virtual character animation generation model to obtain the virtual character animation embeddings corresponding to the animation guiding information after processing the designated prompt information, the animation guiding information and the fully-masked animation sequence by the virtual character animation generation model.

12. The method according to claim 10 or 11, after quantizing and decoding the virtual character animation embeddings to obtain the virtual character action sequence corresponding to the animation guiding information, further comprising:

receiving scene information of a current scene input by a user; and

adjusting the virtual character action sequence based on the scene information using the virtual character animation generation model to obtain an adjusted virtual character action sequence.

13. The method according to claim 10, wherein quantizing and decoding the virtual character animation embeddings to obtain the virtual character action sequence corresponding to the animation guiding information comprises:

receiving target trajectory information input by a user; and

inputting the virtual character animation embeddings and the target trajectory information into a second decoder of a quantized generation model, through decoding processing by the second decoder, to obtain the virtual character action sequence corresponding to the animation guiding information, wherein the quantized generation model is trained based on a plurality of training sample sequences.

14. A data processing system for a task processing model, comprising:

a terminal-side device, arranged for constructing a first sample set and sending the first sample set to a cloud-side device, wherein the first sample set comprises multi-modal sample guiding information; and

the cloud-side device, arranged for inputting the sample guiding information and a sample task sequence into an initial processing model to obtain predicted task embeddings corresponding to the sample guiding information; training the initial processing model according to the predicted task embeddings and sample task embeddings corresponding to the sample task sequence to obtain model parameters of a task processing model in response to reaching a first preset stopping condition, wherein the sample task embeddings are obtained by quantizing and encoding the sample task sequence; and sending the model parameters of the task processing model to the terminal-side device.

15. A computing device, comprising:

a memory and a processor;

wherein the memory is arranged for storing computer executable instructions, the processor is arranged for executing the computer executable instructions, and when executing the computer executable instructions, the processor implements steps as claimed in any one of claims 1 to 13.