process of the model be completed correctly.

Prior to model inference, users need to complete the model initialization through the rkllm_init() function. The specific function definition is as follows:

Table 3-20 Interface Specification for the rkllm_init Function

Fuctiom	rkllm_init
Introduction	Used to initialize the specific parameters and inference settings for RKLLM model.
Parameters	***LLMHandle* handle:*** register model to the corresponding handle for subsequent inference and release calls; ***RKLLMParam* param:*** the parameter structure defined for the model; ***LLMResultCallback callback:*** callback function used to receive and process real-time outputs from the model;
Returns	**0** indicates that the initialization process is normal; **-1** indicates initialization failure;

The example code is as follows:

```
LLMHandle llmHandle = nullptr;
rkllm_init(&llmHandle, &param, callback);
```

3.2.5 Inference Model

After completing the initialization process of the RKLLM model, users can perform model inference using the rkllm_run() function. Real-time inference results can be processed using the callback function predefined during initialization. The specific function definition of rkllm_run() is as follows:

Table 3-21 Interface Specification for the rkllm_run Function

Fuctiom	rkllm_run
Introduction	Used to performing result inference using the initialized RKLLM model.
Parameters	***LLMHandle handle:*** the target handle registered during model initialization. ***RKLLMInput rkllm_input*:*** Input data for model inference. For details, see section