```
'  Qwen'                        IntervitensInc    PaddlePaddle                        ZhipuAI
deepseek-ai    ibm-research    meituan-longcat    neuralmagic    sentence-transformers
wemaster
 AI-ModelScope    JackFram              QuantTrio                        '^Cen'
dengcao              intfloat        meta-llama        openai                shared
'~'
 Anionex              Kwai-Keye        Qwen                                ___root
facebook        jason9693        mistralai        openai-mirror    swift
 ArthurZ              LLM-Research          Shanghai_AI_Laboratory    baichuan-inc
google        jinaai        mlx-community        opendatalab        tclf90
 BAAI              OpenBMB          Tengyunw                cognitivecomputations
hmellor        kernel_meta        models                rhymes-ai        unsloth
 Howeee              OpenGVLab          Yanguan                        damo
home        llava-hf        moonshotai        root                vllm-ascend
root@liteserver-for-vllm-ascend-00004:~/.cache/modelscope/hub/models# cd Qwen
root@liteserver-for-vllm-ascend-00004:~/.cache/modelscope/hub/models/Qwen# ls
QwQ-32B                        Qwen2-VL-7B-Instruct        Qwen2.5-VL-72B-Instruct
Qwen2___5-VL-3B-Instruct        Qwen3-32B                        Qwen3-Reranker-0.6B
QwQ-32B-AWQ                        Qwen2.5-0.5B-Instruct        Qwen2.5-VL-7B-Instruct
Qwen2___5-VL-72B-Instruct                                Qwen3-8B
Qwen3-Reranker-0___6B
Qwen                        Qwen2.5-1.5B-Instruct        Qwen2___5-0___5B-Instruct
Qwen2___5-VL-7B-Instruct                                Qwen3-8B-A3B
Qwen3-VL-235B-A22B-Instruct
Qwen1.5-MoE-A2.7B                Qwen2.5-14B-Instruct        Qwen2___5-14B-Instruct
Qwen2___7-7B                                Qwen3-8B-Base
Qwen3-VL-2B-Instruct
Qwen1___5-MoE-A2___7B        Qwen2.5-32B-Instruct        Qwen2___5-1___5B-Instruct
Qwen3-0.6B                                Qwen3-8B-W8A8
Qwen3-VL-30B-A3B-Instruct
Qwen2-1.5B-Instruct            Qwen2.5-7B                        Qwen2___5-32B-Instruct
Qwen3-0.6B-Base                Qwen3-8B-w4a8                Qwen3-VL-8B-Instruct
Qwen2-1___5B-Instruct            Qwen2.5-7B-Instruct                Qwen2___5-7B
Qwen3-0___6B                Qwen3-Base                        download_model.log
Qwen2-7B                        Qwen2.5-7B-Instruct-1M        Qwen2___5-7B-Instruct
Qwen3-0___6B-Base                Qwen3-Coder-30B-A3B-Instruct    fusion_result.json
Qwen2-7B-Instruct            Qwen2.5-7b-Instruct            Qwen2___5-7B-Instruct-1M
Qwen3-235B-A22B                Qwen3-Embedding-0.6B
Qwen2-7B-W8A8                Qwen2.5-Math-PRM-7B            Qwen2___5-7b-Instruct
Qwen3-235B-A22B-Instruct-2507    Qwen3-Embedding-0___6B
Qwen2-Audio-7B-Instruct    Qwen2.5-Omni-7B                Qwen2___5-Math-PRM-7B
Qwen3-30B-A3B                        Qwen3-Embedding-8B
Qwen2-VL-2B-Instruct                Qwen2.5-VL-32B-Instruct        Qwen2___5-Omni-7B
Qwen3-30B-A3B-Instruct-2507        Qwen3-Next-80B-A3B-Instruct
```

Qwen2-VL-7B   Qwen2.5-VL-3B-Instruct   Qwen2___5-VL-32B-Instruct

Qwen3-30B-A3B-W8A8   Qwen3-Next-A3B-Instruct