

大模型LLM系列课程

原创 唐国梁Tommy 唐国梁Tommy 2024年09月26日 22:06 广东



llama3-01-模型简介

llama3-02-环境配置 模型下载与模型推理

llama3-03-torchrun推理与特殊标记的应用

llama3-04 使用llama.cpp进行llama3模型的量化和部署

llama3-05 使用Docker部署llama3量化模型API服务

llama3-06 基于llama-factory和自定义数据集进行模型微调

llama3-07 基于 TensorRT-LLM 和 Triton 进行 llama3 模型部署-上

llama3-07 基于 TensorRT-LLM 和 Triton 进行 llama3 模型部署-下

Llama 3.2：轻量级AI模型与多模态模型

llama3课程资料请复制以下内容打开百度网盘获取：

链接：<https://pan.baidu.com/s/1xZftTOvHs3GyMx049hIHMA?pwd=5ira> **提取码：5ira**

LoRA_01_算法原理深入解析-上集

LoRA_02_算法原理深入解析-下集

LoRA_03_案例演示 LoRA的核心逻辑

LoRA_04_基于PEFT进行llama3模型微调实战

LoRA_05_01_基于GPT-2进行LoRA微调实战

LoRA_05_02_基于GPT-2进行LoRA微调实战

LoRA_06_LoRA核心算法实现

LoRA_07_在VS Code中debug LoRA微调细节

LoRA课程资料请复制以下内容打开百度网盘获取：

链接：<https://pan.baidu.com/s/13jAb7uo0yiLeQp2xgtC-gA?pwd=wibv> **提取码：wibv**

QLoRA_01_课程简介

QLoRA_02_算法原理_深入解析

QLoRA_03_算法原理_案例实践

QLoRA课程资料请复制以下内容打开百度网盘获取：

链接: <https://pan.baidu.com/s/1zdImsvCWtq8n4t-Aer61pA?pwd=wgmu> **提取码: wgmu**

 点击下方“**资料获取**”可查看全部系列课程

 点击左下角“**阅读原文**”可查看本系列课程全部视频内容

资料获取 16

资料获取 · 目录

上一篇

多模态系列课程

下一篇

LangChain系列课程

阅读原文