

# LLMs 浮点数篇

来自：AiGC面试宝典

宁静致远

2024年01月27日 19:14



扫码  
查看更

- LLMs 浮点数篇
  - 一、fp32和fp16的区别，混合精度的原理
  - 二、半精度是什么？
  - 三、半精度的理论原理是什么？

## 一、fp32和fp16的区别，混合精度的原理

- fp32：32位浮点数，占用4字节，提供高精度的数值表示。
- fp16：16位浮点数，占用2字节，提供相对较低的数值表示精度，但在许多情况下仍足够。
- 混合精度指的是在模型训练中同时使用fp16和fp32，以在计算中提高效率。通常，模型参数使用fp16表示，而梯度和累积值使用fp32表示。这样可以减少内存占用和计算开销，加速训练过程。

## 二、半精度是什么？

半精度是指使用16位二进制浮点数（half-precision floating point）来表示数字的数据类型，可以加速计算和减小内存占用。

## 三、半精度的理论原理是什么？

半精度使用16位二进制浮点数来表示数字，其中1位表示符号位，5位表示指数，10位表示尾数。相比于单精度（32位）和双精度（64位）的浮点数，半精度的表示范围和精度更小，但可以通过降低内存占用和加速计算来实现高效的运算。