

## **1 Introduction**

## **2 Preliminaries**

### **2.1 LoRA**

Among the different techniques for parameter-efficient fine-tuning of LLMs, one of the most relevant approaches of the last years has been LoRA (Low-Rank Adaptation of Large Language Models) [1].

### **2.2 ZKML**

## **3 ZKLoRA**

## **4 Experiments**

## **5 Concluding Remarks**

## **References**

- [1] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021.