

(a) Model Size of Large Language Models (LLM) and Cloud GPU Memory

