GSM8K Performance v.s. Compute cost