**Model Accuracy by Question Type on FinanceBench** Models 1.0 MapReduce gpt-4o-mini Long-Context gpt-4.1-nano Long-Context gpt-4.1-mini 0.84 0.84 0.78 8.0 0.70 0.70 Accuracy 0.0 0.60 0.58 0.52 0.52 0.4 0.2 0.0 domain-relevant metrics-generated novel-generated