

## Cross-Domain Robustness

### In-Domain



Alpaca



BigMath



MMLU

### Out-of-Domain



Magpie



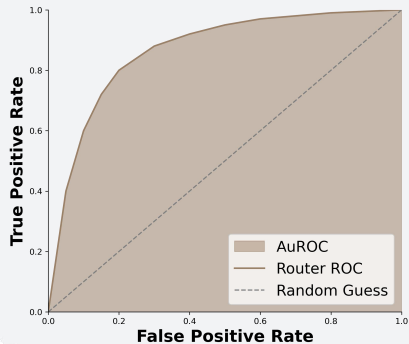
Math



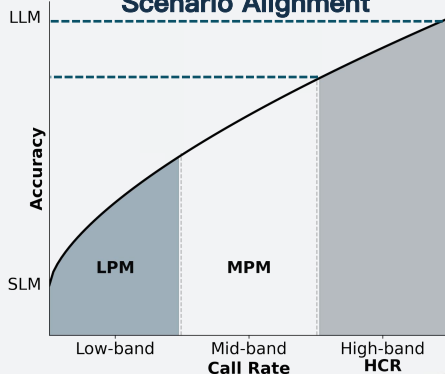
MMLU Pro

## Evaluation Framework

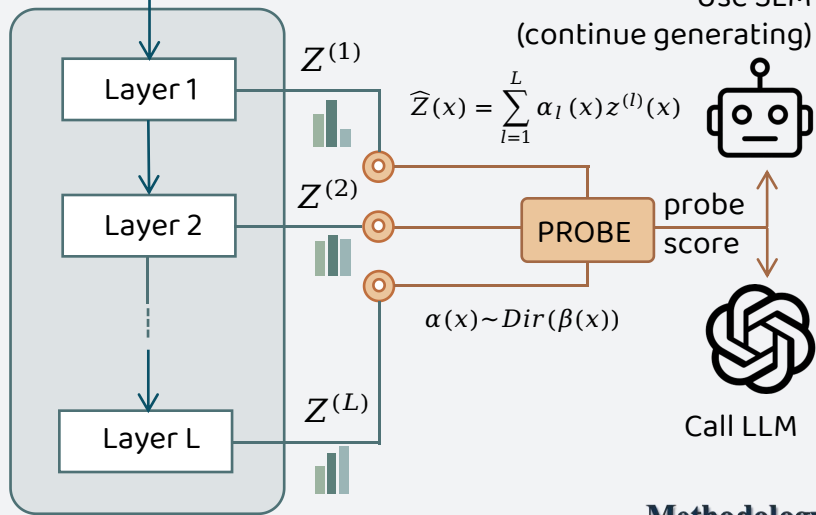
### Router Ability



### Scenario Alignment



prompt



Methodology