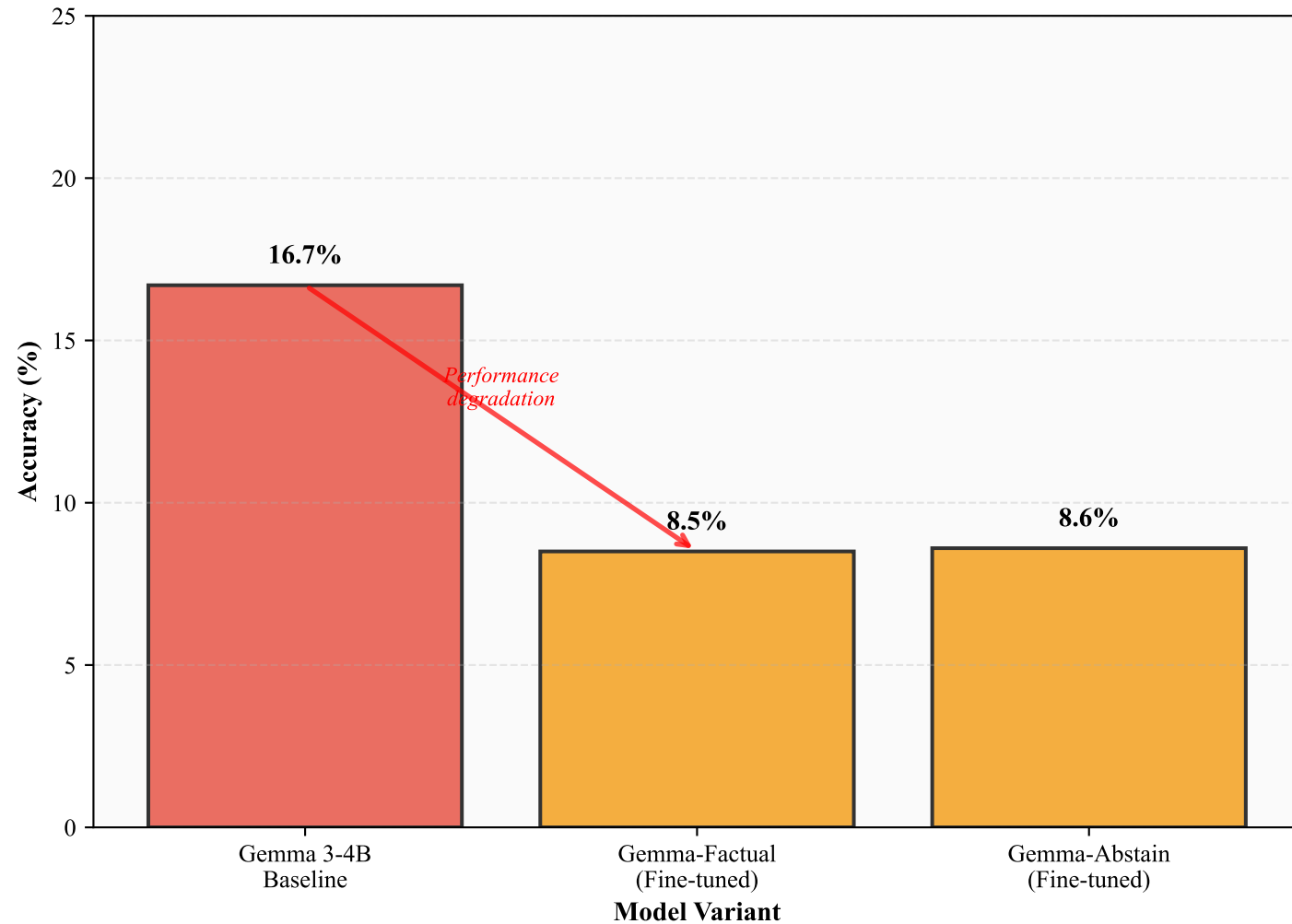


Fine-Tuning Failed to Improve Accuracy



Abstention Behavior: Non-Deterministic
Abstention Precision: 56.7%

