

# Average Build Success Rate Across Reasoning Configurations

## gpt-oss-120b: Low, Medium, and High Reasoning Efforts on DBCode-Bench

