



DeepSeekCoder-1.3B
DeepSeekCoder-6.7B
Qwen2.5-Coder-1.5B
Qwen2.5-Coder-7B

Percentage

Generated Underperforms Human Outperforms Human