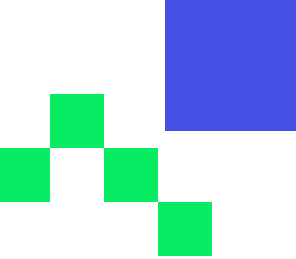




Rogue Deep Think (Preview)

FrontierSWE v2



Preview result
36.78%

Accuracy by verification intensity

