

Spurious SAE Features for GPT-2 at Layer 7

