

# Abliteration Effectiveness: Distribution Preservation vs Refusal Removal

