

GPT-2 Small Pre-trained vs. DART Weights - Sparsity (0%)

Angular Distance

