Accuracy under different training data size and different test difficulty 0.6 0.5 0.4 0.3 0.2 0.1 0.0 D1 D2 D3 Zero-shot with Prompt Few-shot Zero-shot Full data with Prompt Full data Few-shot with Prompt

Accuracy