## D<sub>1</sub>: NLI Sentiment

**context**: When asked about the movie, Ayden said,

'Excellent film. '.

**hypothesis**: Ayden liked the movie

**context**: When asked about the movie, Cannon said, 'Awful. '.

hypothesis: Cannon liked the movie

## <u>D<sub>2</sub></u>: NLI Paraphrase

**context**: someone chronicled that a particular

thing happened
hypothesis: that thing might or might not have happened

**context**: a particular person was satisfied to have a particular thing.

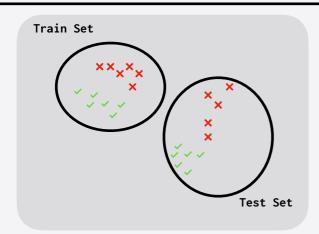
**hypothesis**: that person didn't have that thing

Finetuning Improvement

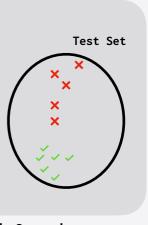
Method	<b>D</b> <sub>1</sub>	D <sub>2</sub>
LMBFF	45.3	-0.4
NullPrompts	43.0	-5.7
BitFit	46.3	-3.5
AdaPET	44.9	-0.3
P-Tuning	46.3	0.3
Few-shot (Avg)	45.2	-2
Full Fine-tuning	45.3	35

(a)

**(b)** 







D<sub>1</sub>: Low Spread, Low Few-shot hardness

 $D_2$ : High Spread, High Few-shot hardness