Task: TFV

Using the given table image, determine if the claim is 'supports' or 'refutes'. If the table does not provide sufficient information, the answer should be 'not enough info'. Output the final answer as JSON in the format {'answer': '<YOUR ANSWER>'}. BERT shows a gain in accuracy for the QNLI target corpus.

Model	Target Corpus		Transfer Filler	Transfer Role	Baseline Acc. (%)		Gain (%)
BERT	QNLI	True	_	_	91.60	91.27	_ 0.33
BERT	QQP	True	_	_	91.45	91.12	_ 0.33
BERT	RTE	True	_	_	71.12	73.65	+ 2.53
BERT	SNLI	True	_	_	90.45	90.69	+ 0.24
BERT	SST	True	_	_	93.23	92.78	_ 0.45
HUBERT (Transformer)	QNLI	True	True	False	90.56	91.16	+ 0.60
HUBERT (Transformer)	QQP	False	False	True	90.81	91.42	+ 0.61
HUBERT (Transformer)	RTE	True	True	True	61.73	74.01	+ 12.28
HUBERT (Transformer)	SNLI	True	False	True	90.66	91.36	+ 0.70
HUBERT (Transformer)	SST	True	False	True	91.28	92.43	+ 1.15

Table 3: Transfer learning results for GLUE tasks. The source corpus is MNLI. Baseline accuracy is when Transfer BERT, Filler, and Role are all False, equivalent to no transfer. Fine-tuned accuracy is the best accuracy among all possible transfer options.

Output: correct, wrong Answer: Refutes

`json\n{\n \"answer\": \"supports\"\n}\n```



Ours (Qwen2-VL-7B-Ins



The table shows that for the QNLI target corpus, the gain for BERT is -0.33%, indicating a decrease in accuracy. Thus, the answer is {\"answer\": \"refutes\"}.

Intern2-VL-8B



The answer can be represented in JSON as follows: $\n\i''json\n {\n ''answer'': \n'''n}\n'''}$





Based on the table information, the given claim is false. Thus, the answer is {\"answer\": \"refutes\"}.

LLaVA-NeXT-7B



The statement is also ambiguous as it refers to \"QNLI target corpus,\" which could refer to a specific dataset or context, but the image does not provide enough information to understand this context.