Task	Table Image	Input: Task, Output format, Task-related input	Output: Reasoning, Answer or description

Task | **Table Image** | **Input:** Task, Output format, Task-related input | **Output:** Reasoning, Answer or description

Table Question Answer

Model	Target Corpus	Transfer BERT	Transfer Filler	Transfer Role	Baseline Acc. (%)	Fine-tuned 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.

Input (Table Question Answer):
Refer to the provided table image and work through the question step by step. Output the final answer as JSON in the format {'answer': '<YOUR ANSWER>'}. Question: 'Which model shows the highest gain in accuracy for the RTE target corpus

Output (Table Question Answer):
By comparing the gain percentages for the RTE target corpus, HUBERT (Transformer) shows a gain of +12.28%, which is higher than BERT's gain of +2.53%.", **So, the answer is {"answer": "HUBERT (Transformer)"}.**

Table Fact Verification

Input (Table Fact Verification):
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.

Output (Table Fact Verification):
The table shows that BERT has a gain of -0.33% for the QNLI target corpus, indicating an decrease in accuracy. **Thus, the answer is {\"answer\": \"refutes\"}.**

Table to Text Generation

Input (Table to Text Generation):
You are provided with a table image with caption. Formulate a statement describing the table.
Table Caption: '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 (Table to Text Generation):
HUBERT (Transformer) shows substantial gains across all target corpora after transfer learning from MNLI, while BERT shows a drop in accuracy for QNLI, QQP, and SST. Table 3 summarizes the results for these transfer learning experiments when the source task is MNLI. Gain shows the difference between Fine-tuned model's accuracy and Baseline's accuracy. For HUBERT (Transformer), we observe substantial gain across all 5 target corpora after transfer. However, for BERT we have a drop for QNLI, QQP, and SST.