

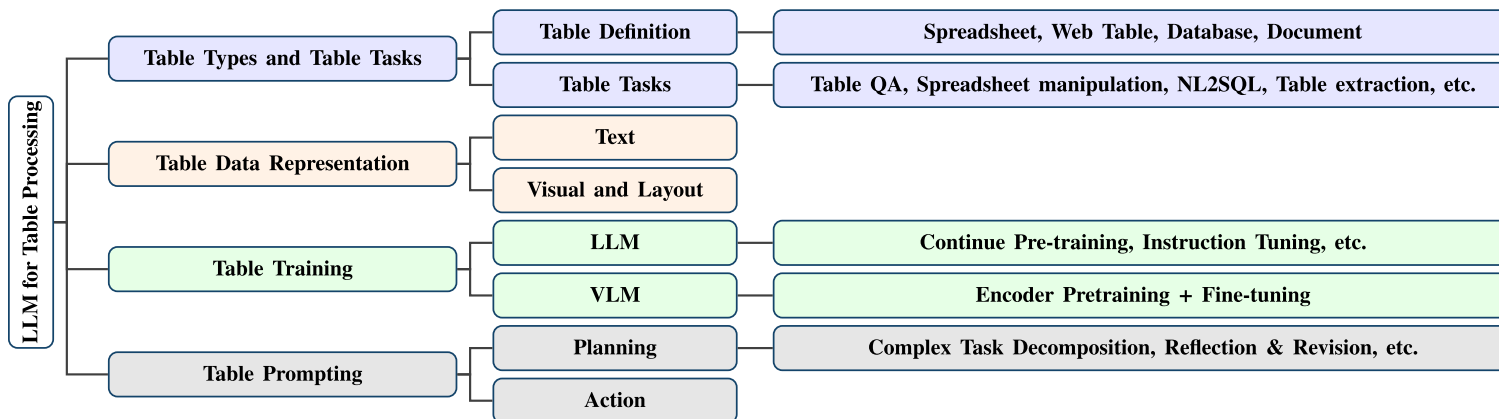
Large Language Model for Table Processing: A Survey

Weizheng LU, Jing ZHANG, Ju FAN, Zihao FU,
Yueguo CHEN, Xiaoyong DU

Frontiers of Computer Science, DOI: [10.1007/s11704-024-40763-6](https://doi.org/10.1007/s11704-024-40763-6)

Background & Ideas

- Background of **Large Language Model for Table Processing**:
 - Large volumes of data stored in tables: Spreadsheet, Web Table, Database, Document
 - Industry and academia want to use Large Language Models (LLMs) or Visual Language Models (VLMs) to automate table processing
- Ideas: review latest developments in LLM for table processing, summarize table tasks and types, categorize methods, collect datasets and benchmarks.



Main Contributions

- Contributions:
 - cover diverse **table tasks**;
 - categorize **methods**: training and prompting;
 - instruction training
 - chain-of-thought, ReAct
 - discuss **challenges**.

An example entry from TableInstruct

Instruction:

This is a table fact verification task. The goal of this task is to distinguish whether the given statement is entailed or refuted by the given table.

Table:

[TLE] The table caption is about tony lema. [TAB] | tournament | wins | top - 5 | top - 10 | top - 25 | events | cuts made [SEP] | masters tournament | 0 | 1 | 2 | 4 | 4 | 4 | [SEP] ...

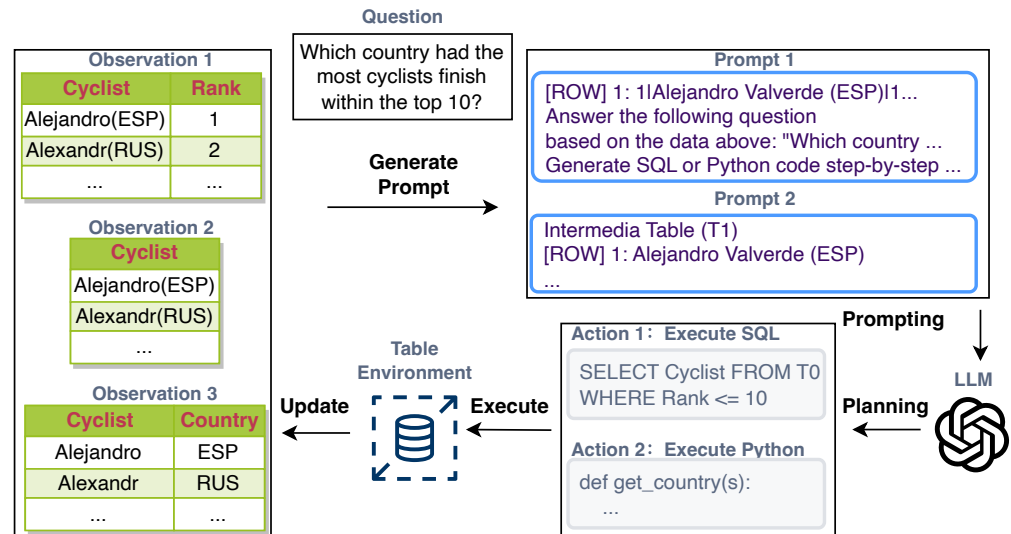
Question:

The statement is:<tony lema be in the top 5 for the master tournament, the us open, and the open championship>. Is it entailed or refuted by the table above?

Output:

Entailed.

Training:
Build Instruct Tuning Dataset



Prompting
Agent Workflow