NormTab: Improving Symbolic Reasoning in LLMs Through Tabular Data Normalization

Md Mahadi Hasan Nahid, Davood Rafiei -> {mnahid, drafiei}@ualberta.ca

Department of Computing Science, University of Alberta, Canada

1. Overview

We introduce **NormTab**, a framework that uses LLMs to normalize web tables into structured relational database formats, improving symbolic reasoning tasks.

2. RQs and Key Features

Research Questions:

- **RQ1:** How can we leverage LLMs' textual understanding to effectively clean and normalize web tables?
- **RQ2:** How can web table normalization enhance table reasoning tasks, particularly in the context of LLM based symbolic reasoning?

Key Features:

- **(1)Structure normalization:** transposing tables, flattening rows and columns.
- **(2) Value normalization:** removing extraneous strings, standardizing dates and numbers.
- **(3)** Effective for data cleaning and transformation tasks.
- **(4)** Addresses structural variance, mixed values, noise, and substring extraction in web tables

4. Experimental Setup

LLM: gpt-3.5-turbo-0125, gpt-4-turbo, gemini-1.5-flash
Datasets: **(1)** WikiTableQuestions **(2)** TabFact
Evaluation Process:

- We apply Text-to-SQL on Original tables and Normalized tables
- Compare the performance

6. NormTab Accuracy

Type	Accuracy
Columns Selection	91.0%
Transpose Detection	97.0%
Last Row Aggregation	100.0%
Split Column	87.0%
Date and Number	100.0%
N/A value	93.0%
Value Cleaning	82.0%

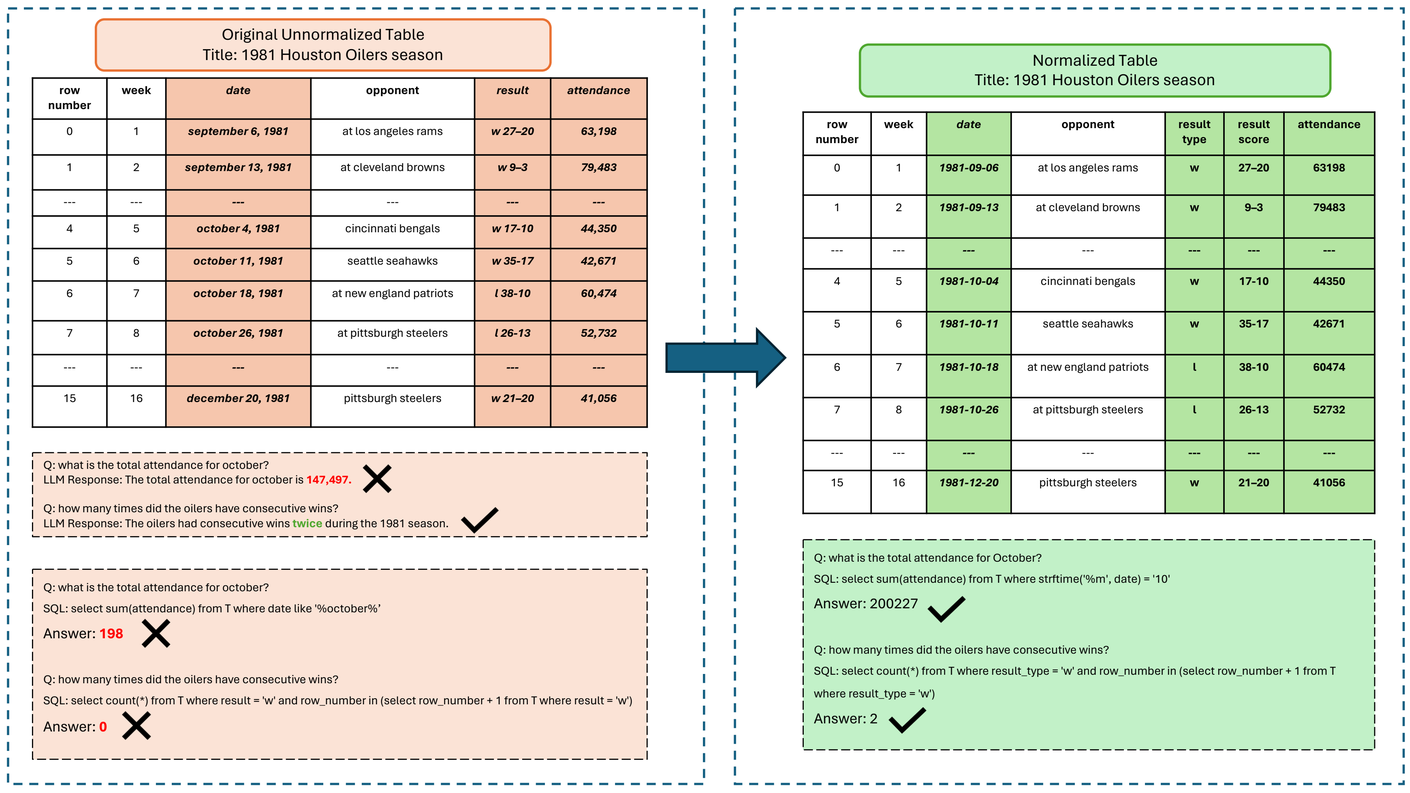
Table 4: Accuracy of ***NormTab*** in various types of normalization.

7. NormTab: Basic vs Targeted

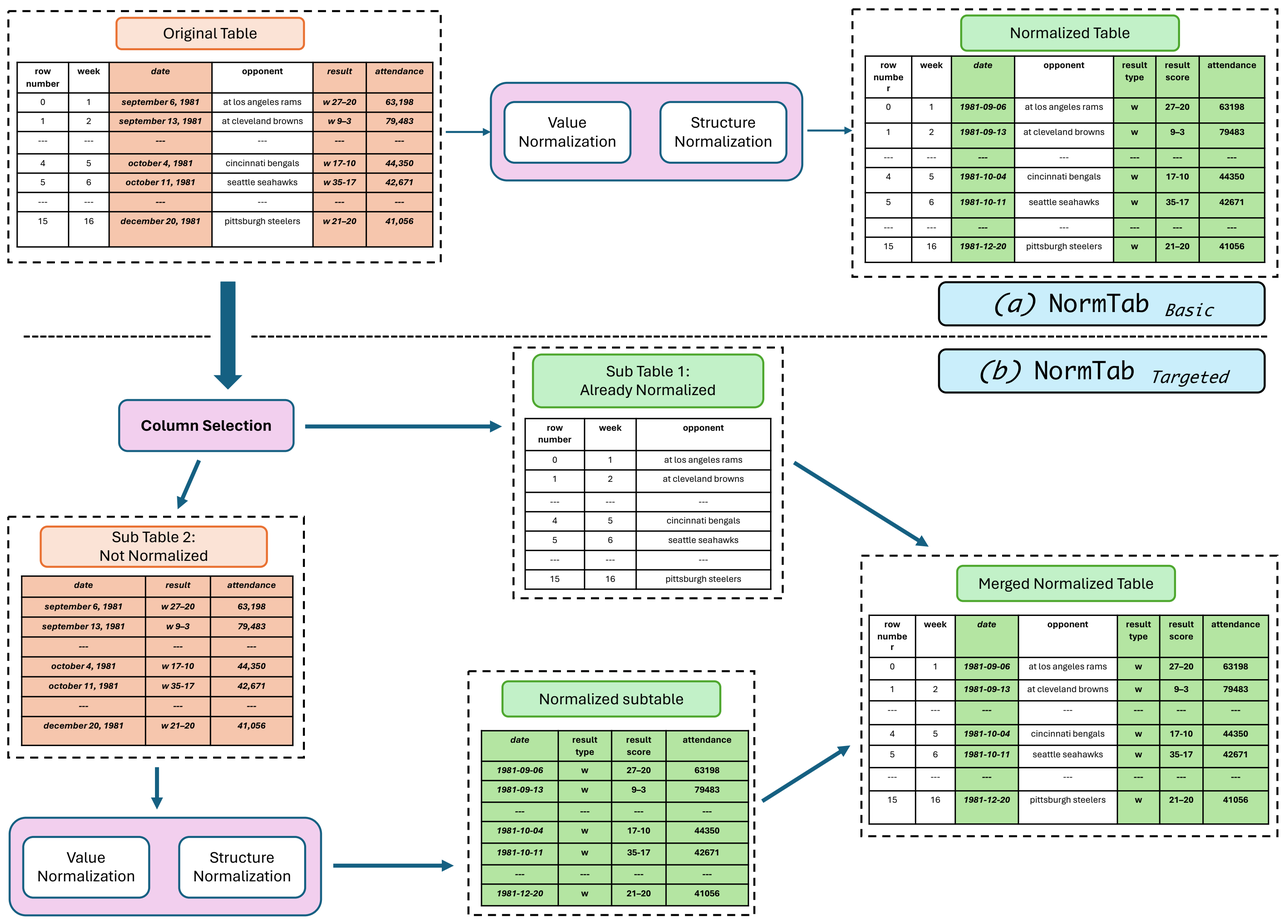
Dataset	NormTab (Basic)	NormTab (Targeted)	Reduction
WikiTQ	152.26	41.82	**72.53%**
TabFact	106.19	29.11	**72.58%**

Table 8: Table cell reduction in NormTab-Targeted compared to NormTab-Basic

3A. Motivation



3B. Methodology



5. Results

Model	Acc (%)
TableCoT (Chen, 2023)	52.40
BINDER	56.74
DATER	52.80
ReAcTable	52.40
Rethinking-Tab-Data	56.87
Chain-of-Table	59.94
Text-to-SQL (Rajkumar et al., 2022)	52.90
Text-to-SQL (gpt-3.5-turbo)	51.30
NormTab$_{Basic}$ + SQL (ours)	**60.80**
NormTab$_{Targeted}$ + SQL (ours)	**61.20**

Table 1: Performance comparison of NormTab on WikiTQ dataset. The results clearly demonstrate that NormTab significantly surpasses other models in accuracy when employing symbolic reasoning.

Model	Acc (%)
TableCoT-chatgpt	73.10
BINDER	79.17
DATER	78.01
Chain-of-Table	80.20
ReAcTable	73.10
Text-to-SQL (Rajkumar et al., 2022)	64.71
Text-to-SQL (gpt-3.5-turbo)	62.32
NormTab$_{Basic}$ + SQL (ours)	**67.10**
NormTab$_{Targeted}$ + SQL (ours)	**68.90**

Table 2: Performance comparison of NormTab on TabFact dataset with other models.

Model	Acc (WTQ)
NormTab$_{Basic}$ (gemini-1.5-flash)	61.36
NormTab$_{Targeted}$ (gemini-1.5-flash)	61.24
NormTab$_{Basic}$ (gpt-4-turbo)	61.57
NormTab$_{Targeted}$ (gpt-4-turbo)	**62.28**

Table 3: Performance of NormTab on WikiTQ dataset using Gemini-1.5-flash and GPT-4-turbo.

Model	Acc (WTQ)
TabSQLify	64.7
NormTab+TabSQLify	**68.63**

Table 5: Performance of TabSQLify integrated with NormTab on the WikiTableQuestions dataset (gpt-3.5-turbo).

8. Conclusions

Here are some **key takeaways**:

1. Utilizes LLMs' text comprehension for data cleaning and normalization.
2. Enhance symbolic code generation performance.
3. Demonstrates significant improvement in symbolic table reasoning.