

# IndoMPF: Dataset and Tabular-Graph Numerical Reasoning Baseline Model for Generating Notes to Government Financial Statements

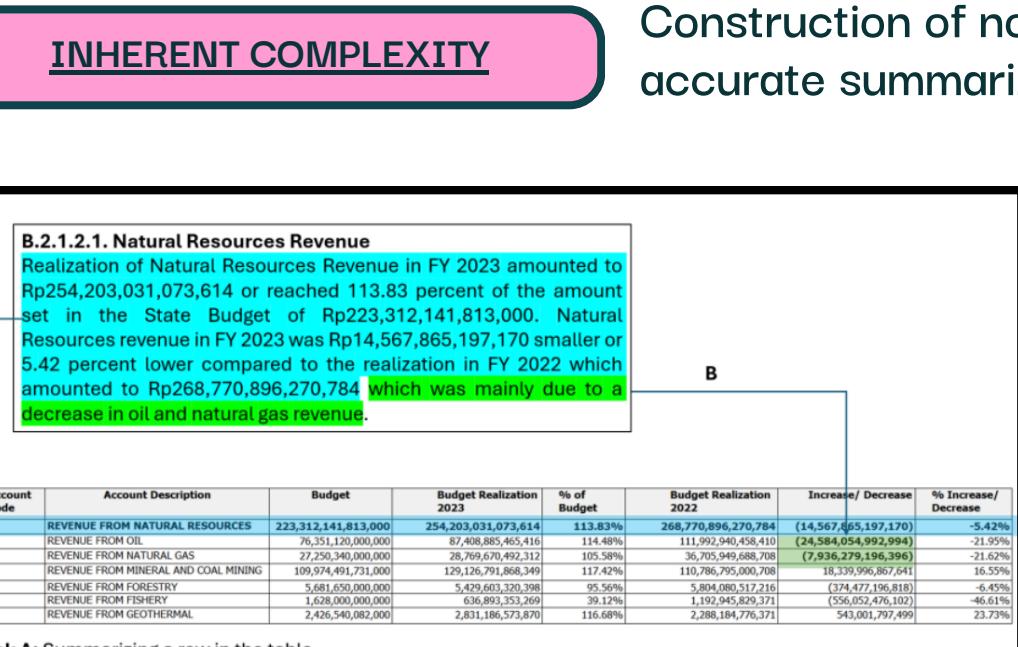
## BACKGROUND



Notes to financial statements accompanies a financial statement to enhance its understandability and to avoid misunderstanding.



The current development of Language Model that allows for various text-generation tasks brings the potential of automating Notes to Financial Statement generation.



**B.2.1.2.1. Natural Resources Revenue**  
 Realization of Natural Resources Revenue in FY 2023 amounted to Rp223,312,141,813,000. The realization of Central Government Expenditure in the State Budget of Rp223,142,465,193,170 amounted to Rp2,312,141,813,000. Natural Resources Revenue in FY 2023 was Rp14,567,465,193,170 more than the realization of expenditure in the State Budget of Rp223,142,465,193,170. The total realization of expenditure in the State Budget of Rp223,142,465,193,170 was 15.13% higher than the realization of Natural Resources Revenue in FY 2023 which amounted to Rp223,312,141,813,000.

**Central Government Operating or Organization Budget Division**  
 The realization of Central Government Expenditure in FY 2023 according to the Budget Section was the largest in 1999 (State Treasury) amounting to Rp1,267,557,264,455,000. The realization of Central Government Expenditure in the State Budget of Rp223,142,465,193,170, the total realization of Central Government Expenditure in FY 2023 was Rp1,267,557,264,455,000. The realization of expenditure in the State Budget of Rp223,142,465,193,170 was 15.13% higher than the realization of Natural Resources Revenue in FY 2023 which amounted to Rp223,312,141,813,000.

**To calculate the 15.13 percentage increase, the expenditure from state treasury must be excluded from total expenditure**

## OBJECTIVE

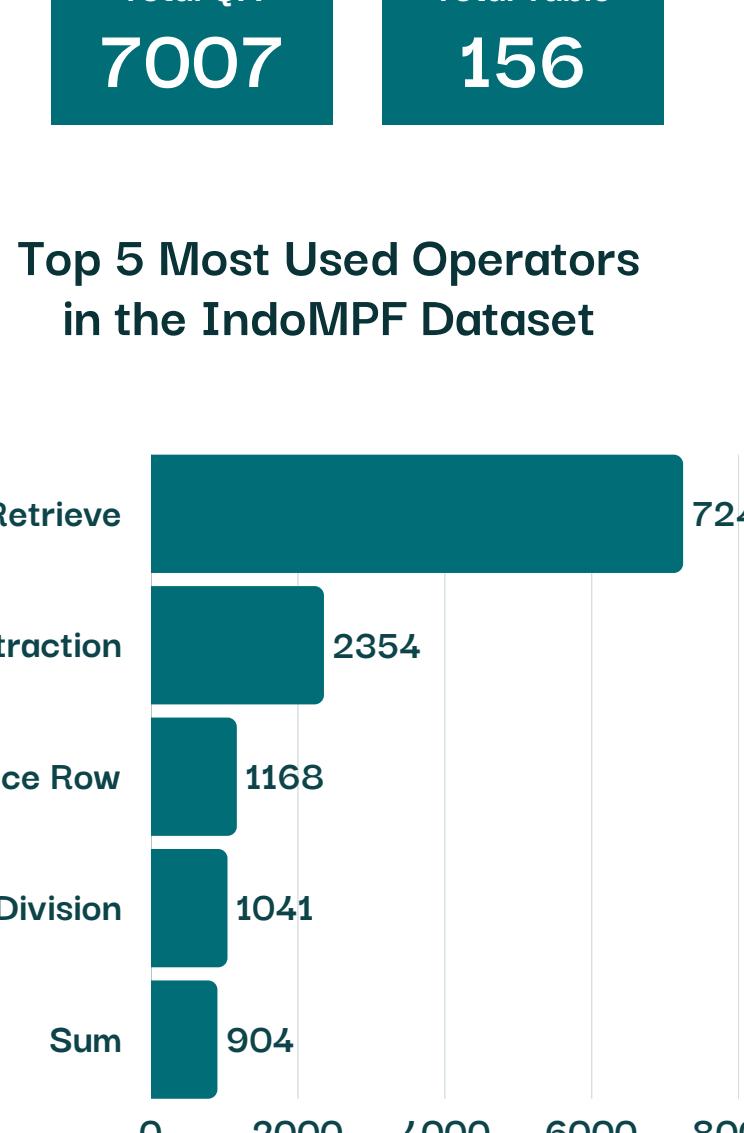
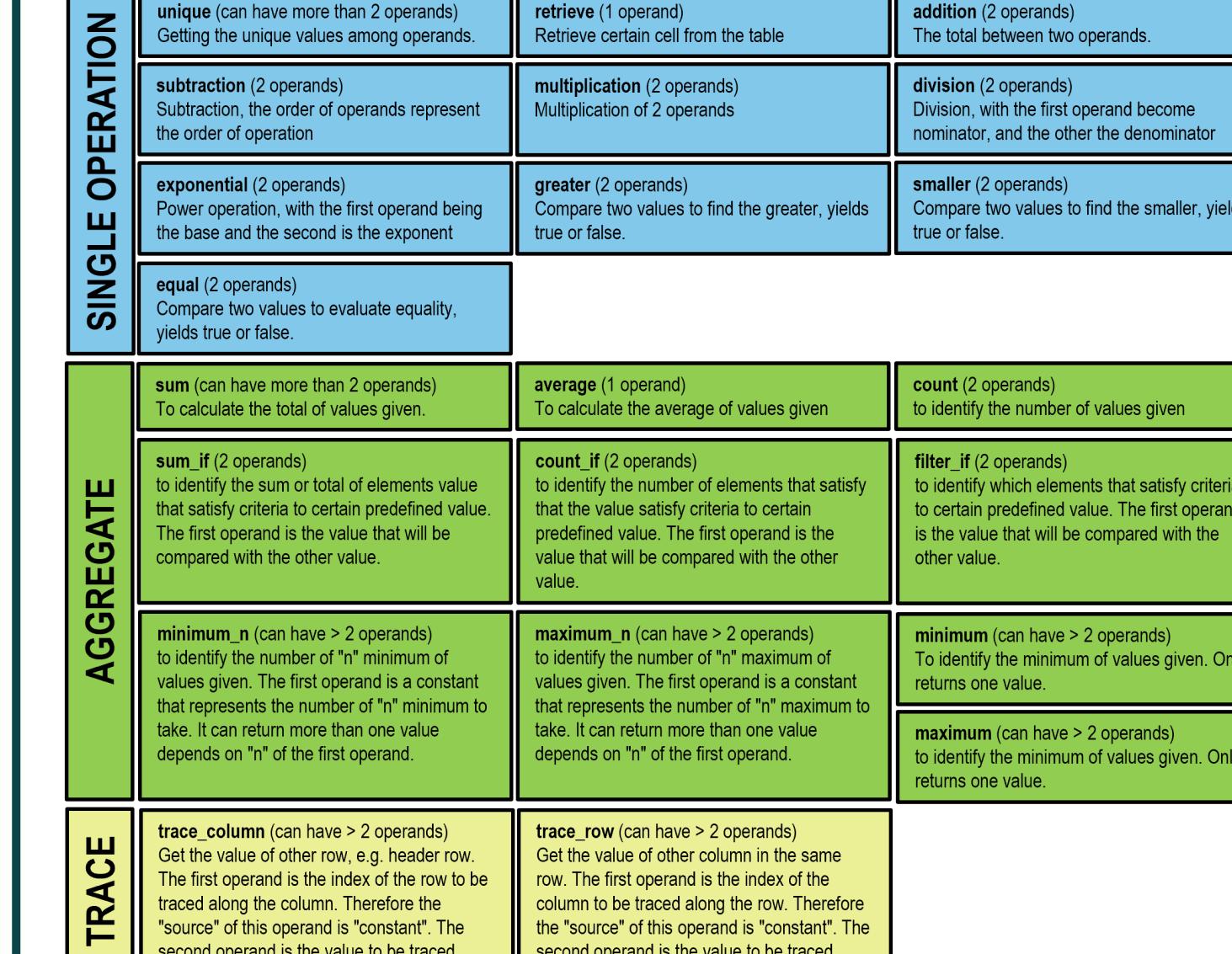
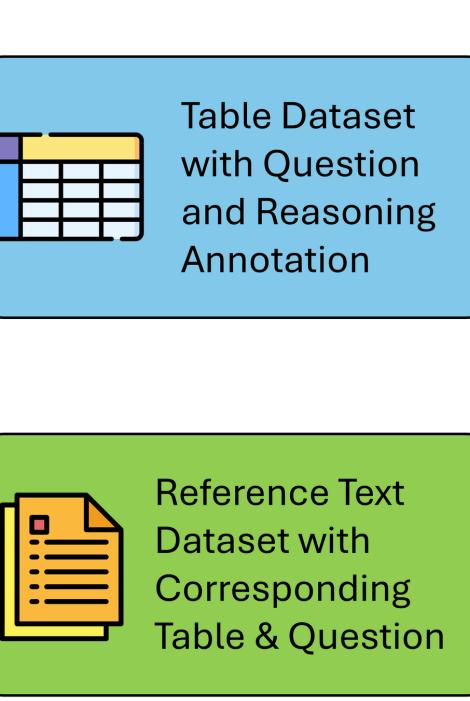
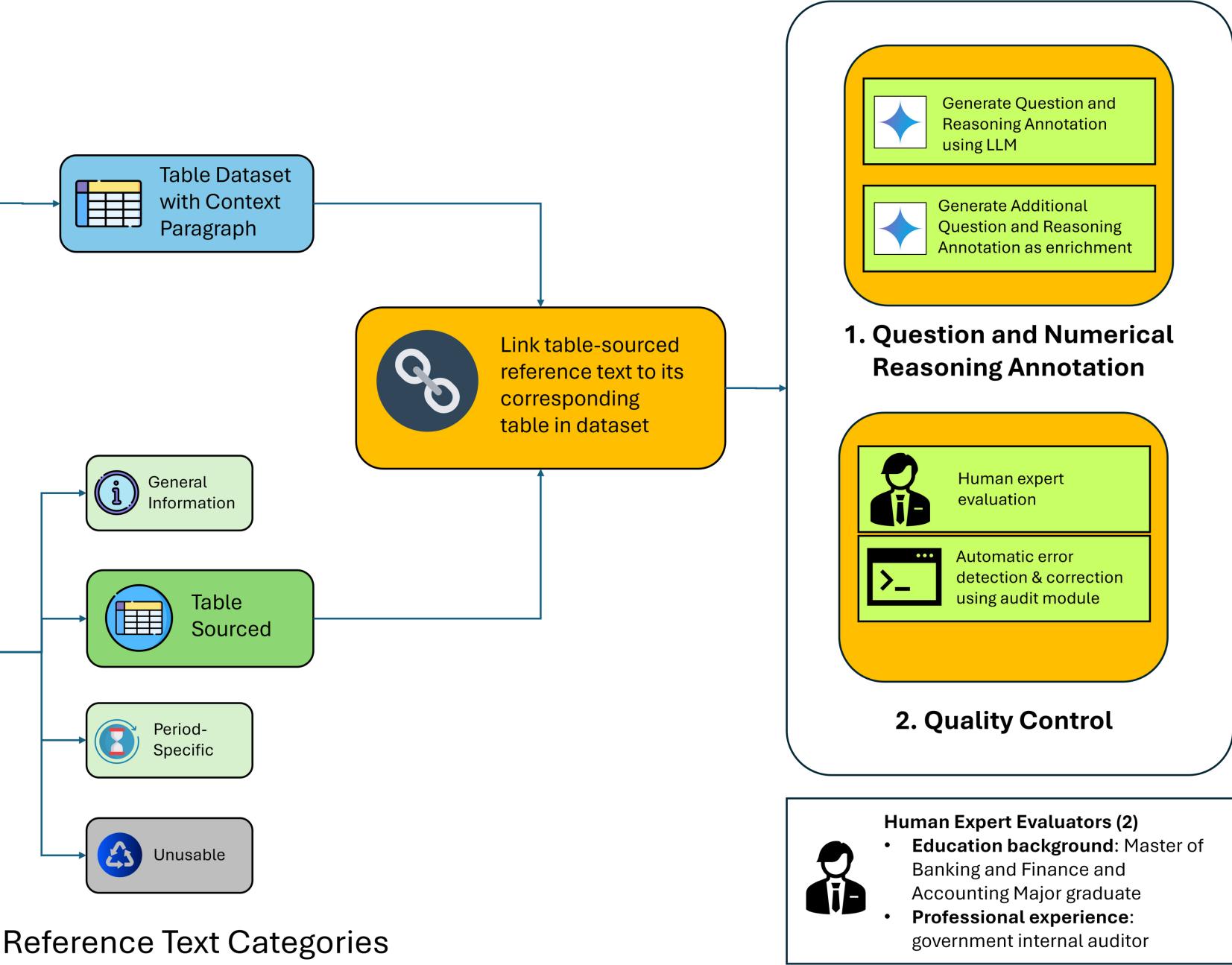
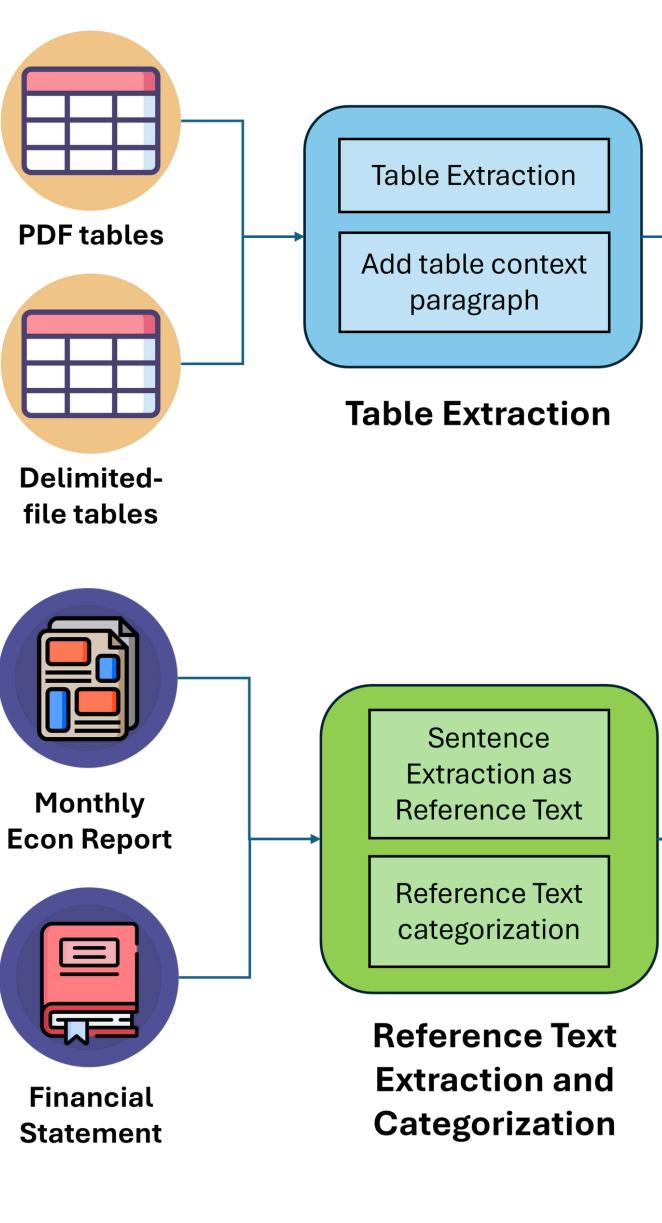
Construction of table question answering dataset with numerical reasoning with specialized domain in macroeconomic and public finance

Fine tuning of embedding model using constructed domain-specific dataset for table retrieval task.

Development of baseline model for table question-answering with numerical reasoning

## DATASET CONSTRUCTION

### Dataset Sources



## RETRIEVAL MODEL

### EMBEDDING MODEL FINETUNING

Base embedding model:



Finetuning Dataset

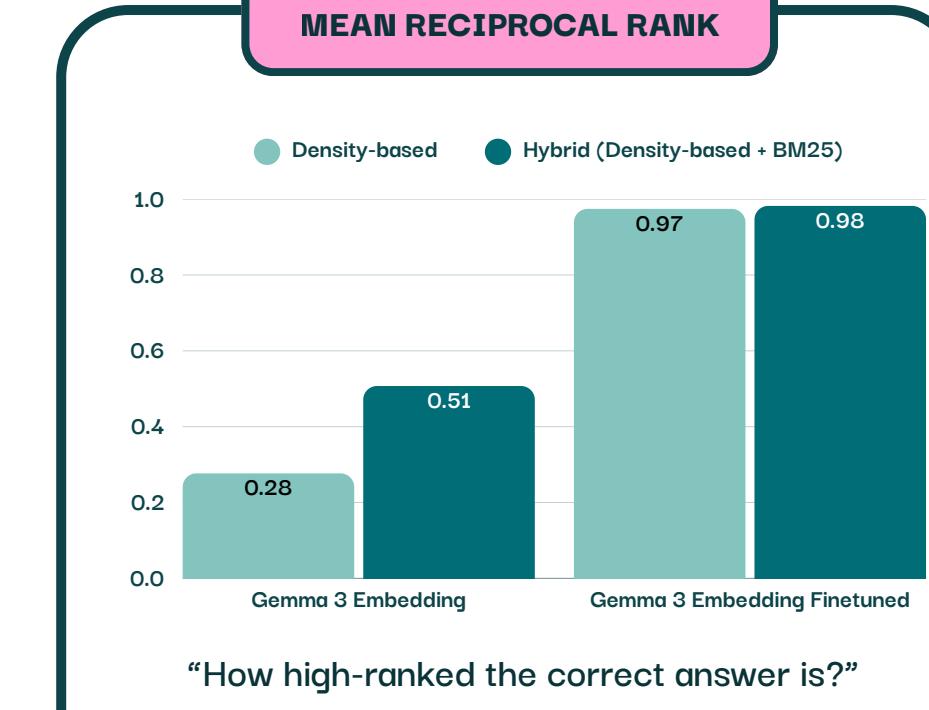
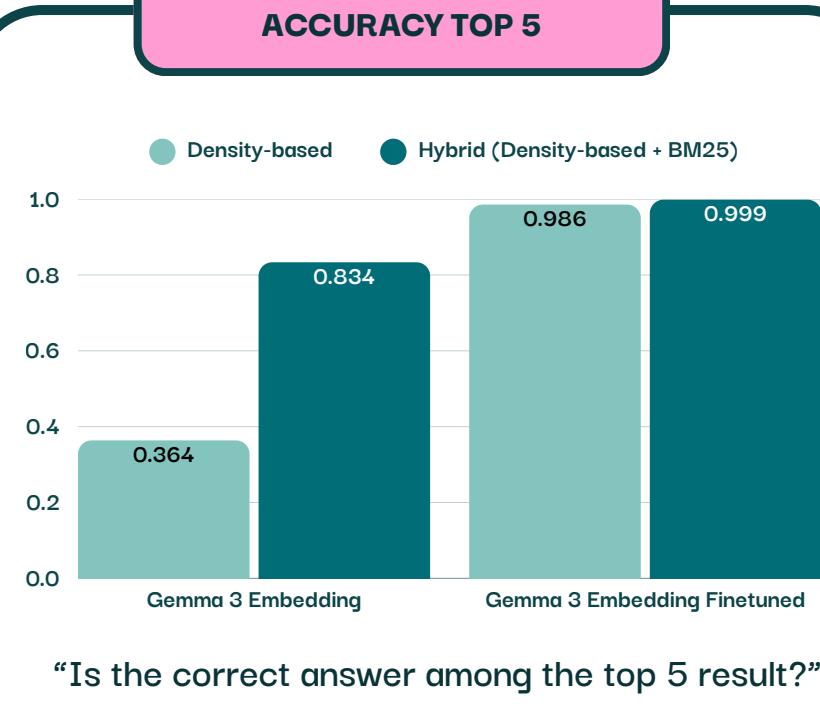
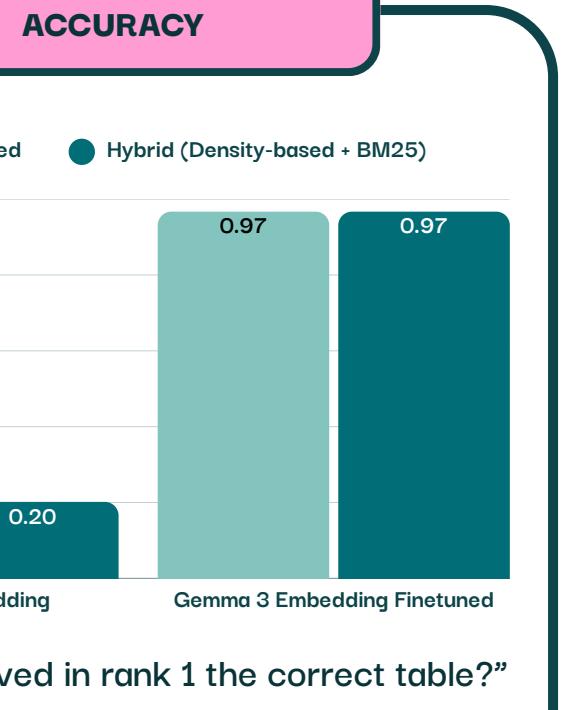
Table Dataset with Question and Reasoning Annotation

Finetuning Strategy

### Multiple Negative Ranking

Contrastive learning technique to train embedding model to distinguish relevant and irrelevant text.

Which table can be used to answer a question?



- Finetuning was the Primary Success Factor: The use of Multiple Negatives Ranking (MNR) loss for embedding model finetuning objective was exceptionally effective, leading to a nearly five-fold increase in top-1 accuracy. This confirms that for specialized retrieval tasks, finetuning is not just beneficial—it is essential for achieving state-of-the-art performance.
- Hybrid Retrieval is a Powerful Secondary Enhancement: The addition of BM25 provides a crucial boost to recall, ensuring that even if the dense model misses, the keyword-based search can often find the correct document. It is a low-cost, high-reward addition to the system.
- The Combined System is Production-Ready: The final strategy, combining the MNR-finetuned embedding model with a hybrid (dense + BM25) retrieval system, has achieved near-perfect results on the test set. With an Accuracy@1 of 96.7% and Accuracy@5 of 99.9%, this system is highly reliable and can be confidently deployed. No further finetuning is recommended at this stage, as the performance metrics are already approaching their practical ceiling.

## INDOMPFGATE: BASELINE TABLE QUESTION-ANSWERING MODEL

### Graph Representation of Input

Table, question, and explanatory paragraph are represented as graph, to better represent structural information of the table and effectively capture the relation between table and texts.

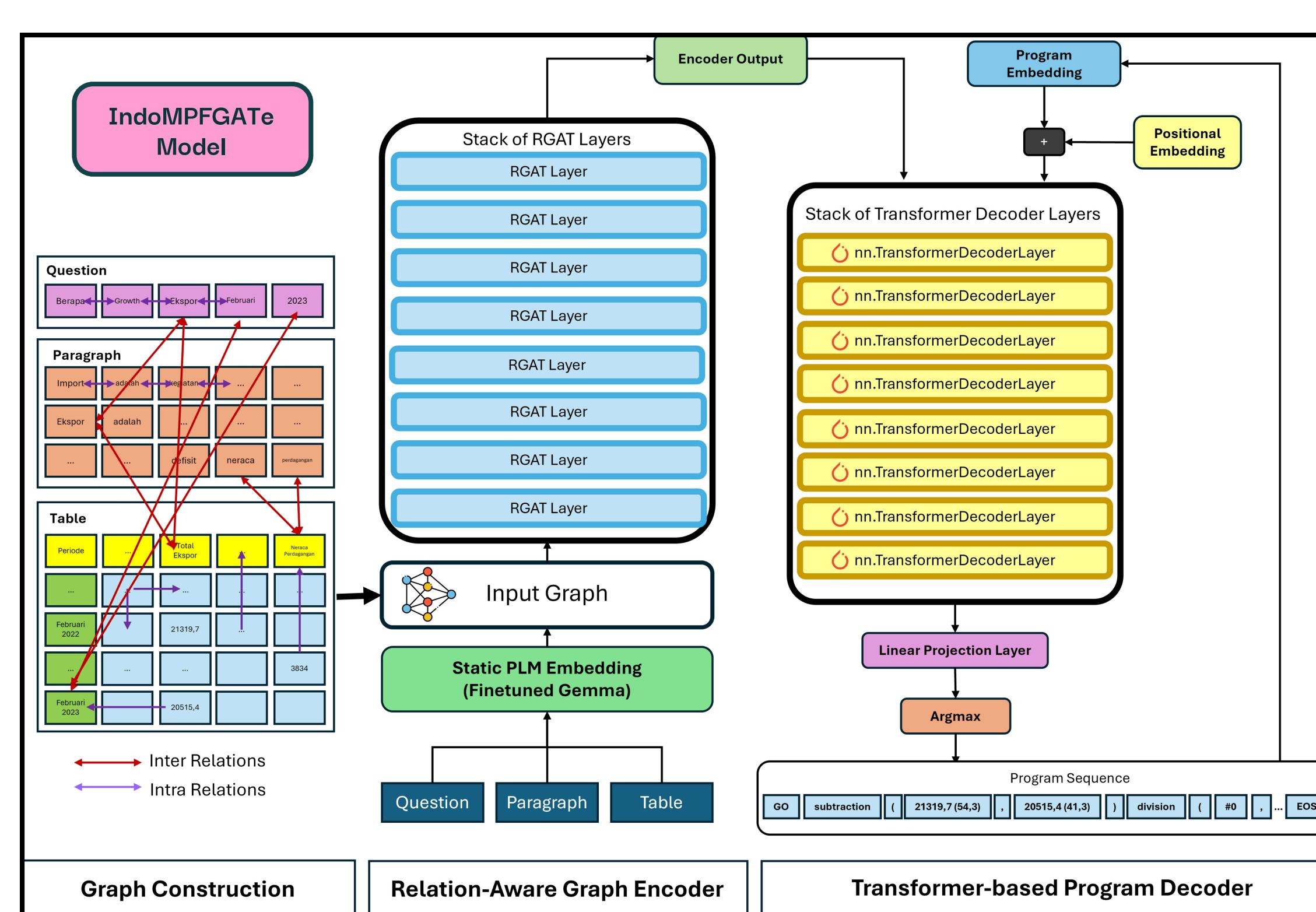
Source x	Source y	Relation	Description
<b>Intra-Relation</b>			
Qword	Pword	question to next word	y is the next word of x in the question.
Pword	Pword	paragraph to next word	y is the next word of x in the paragraph.
Pword	Pword	paragraph to paragraph word	y is the word in the paragraph x.
Tcell	Tcell	table cell to next cell in row	x is adjacent to cell y in the same table row.
Tcell	Tcell	table cell to next cell in col	x is adjacent to cell y in the same table column.
Tcell	Tcell	table cell in the same row	x and y are cells in the same table row.
Tcell	Tcell	table cell in the same col	x and y are cells in the same table column.
Tcell	Tcell	table cell to row header	x is in a row where cell y is the row header.
Tcell	Tcell	table cell to col header	x is in a column where cell y is the column header.
<b>Inter-Relation</b>			
Qword	Pword	question to paragraph word	x and y is the common word of question & paragraph.
Qword	Pword	question to table	Table cell x contains word x in the question.
Tcell	Pword	table to paragraph word	Paragraph y contains cell x.
Tcell	Pword	table to table	Table cell y contains paragraph word y.

### Graph Encoder

Aggregates information about nodes and edges of graph. Relation awareness is incorporated through embedding of edges.

### Transformer Decoder

Autoregressive decoder that generates the program token by token. During prediction, masking is applied to ensure program generated is grammatically correct.



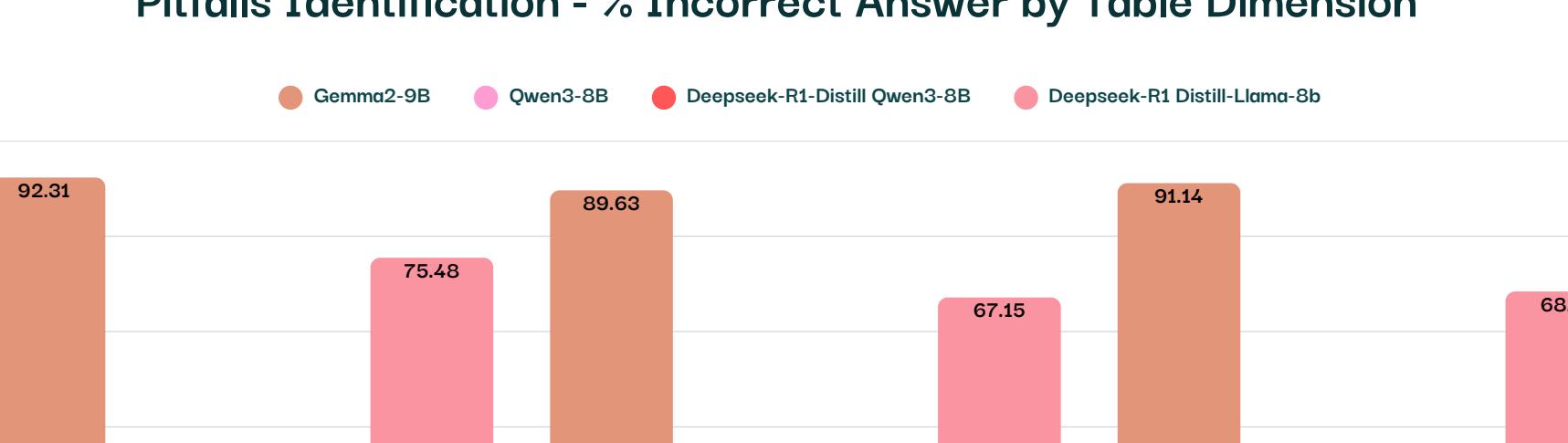
### Results

Epoch 85

Validation F1: 0.0045  
Validation EM: 0.0038  
Test F1: 0.0045  
Test EM: 0.0038  
Can only answer 1 question correct.

## FUTURE WORKS

### Pitfalls Identification - % Incorrect Answer by Table Dimension



### Pitfalls Identification - % Incorrect Answer by Reasoning Step Length

