Graph Ingestion Engine (Fall 2024)

Mohamed Essam Moustafa Kamel Fayed, Bruce Walker
mohamed.fayed@gatech.edu, bruce.walker@psych.gatech.edu

***Abstract*—**Multimodal Large Language Models have shown impressive visual capabilities in many Visual Question Answering tasks. In this paper, we aim to test them on Chart-to-Table task.

CONTENTS

Extracting data points from charts has witnessed attention from research community. There has been work to summarize charts content end-to-end. Another direction was related to converting charts to tables. Recently, there has been efforts to analyze the performance of Large Visual e Models (LVMs) in many all of those tasks. In our work, we aim to pay closer attention to Chart-to-Table task.

- Gemini 1.5 Flash
- ChartGemma
- Deplot

- Relative Mapping Similarity (RMS)

For the sample we analyzed, there has been no errors in recognizing text in the images, e.g. columns names. Moreover, the tables layouts were perfectly generated into table in json format.

4.2 Values Extraction

For PlotQA and ICPR22 samples, it is frequent to find errors like:

1. rounding errors, e.g. $15.42- > 15 and 15.6- > 15$.
2. Precision Errors: we have noticed that the model can not predict more than 3 digits for each value, e.g. $126765000.0- > 156000000$.
3. In case of near values, e.g. 24, 24, there might be some errors, e.g. predicting 23 instead of 24. For that kind of error, it may result in changing trend, e.g. steady performance may seam as decreasing. [1]
4. There are some graphs, like **??**, the API just fails with no clear response message (till now). However, it is suspected that the very large number of data points might be the reason.

It is also worth noting that the model can differntiate outputs based on scale, e.g. 156000000&50.2 for instance.

Table **??** includes Gemini 1.5 Flash predictions for figure **??**

Table 1—Reference Data for PlotQA Vertical Bar Chart # 25905

Country	2005	2006	2007
Belarus	31.25	24.359312083694	24.3710602532796
Belize	15.4217891420299	15.6256596494276	13.8130345743766
Bosnia and Herzegovina	33.3030438301411	34.2434876503054	50.3046671142766
Brazil	44.5074830208778	47.9767970144999	52.58720265358
Cabo Verde	11.0244112659453	7.615772874299 68	7.81923662180111
Canada	51.80682953179811	51.9264969983905	51.59612102301951

5 CONCLUSION AND RECOMMENDATIONS

In this report, we document our quanitative analysis for LLMs behavior in Chart-to-Table task. Based on the selected sample, we observed that the model can accurately recognize the layout of the graph, but it is not very precise in recognizing small differences in values. For future work, we recommend combining both

1 It is worth noting that we have not seen cases where increasing is replaced by decreasing trends or vice versa.

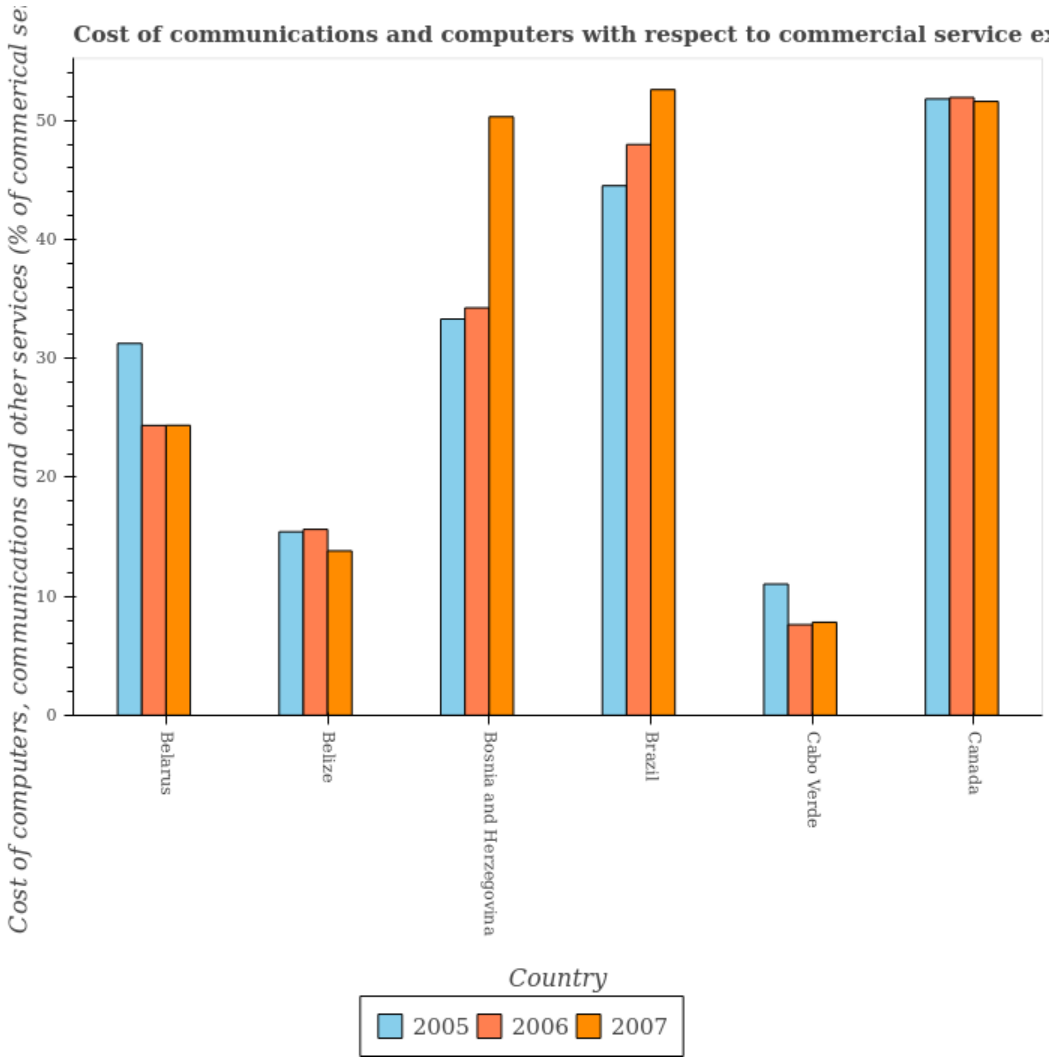

Figure 1—Vertical Bar Chart example from PlotQA testset.

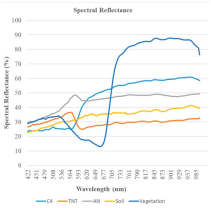

Figure 2—Example for charts that causes the API to fail.

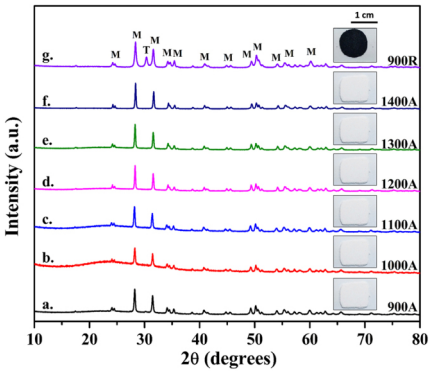

Figure 3—Another example for charts that causes the API to fail. It is also a good example of bad graphs in the wild.

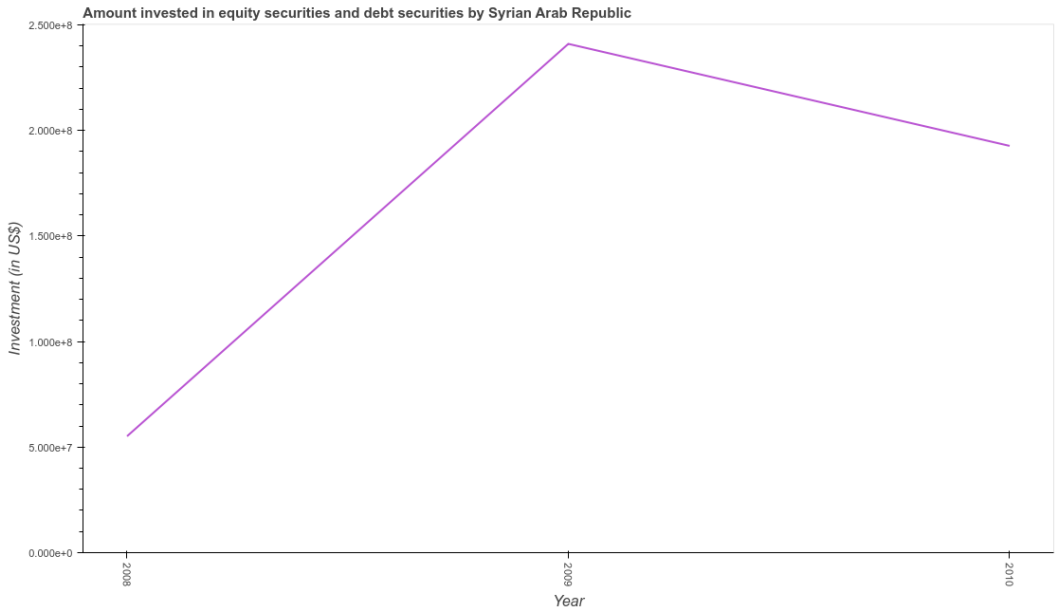

Figure 4—Example for Line Chart from PlotQA testset # 21673 about Portfolio Investment

Table 2—Gemini 1.5 Flash Prediction Data for PlotQA Vertical Bar Chart # 25905

Country	2005	2006	2007
Belarus	31.0	24.0	23.0
Belize	15.0	15.0	13.0
Bosnia and Herzegovina	33.0	34.0	51.0
Brazil	44.0	47.0	52.0
Cabo Verde	10.0	7.0	8.0
Canada	51.0	51.0	50.0

Table 3—Gemini 1.5 Flash predictions for PlotQA vertical Bar #25905

x	y
0	55132083.68
1	241005430.1
2	192682327.3

Table 4—Reference for Line Chart from PlotQA #21673 Portfolio Investment

LLMs and Computer Vision algorithms to complement each other in accurately converting charts into tables. [2]

Year	Investment (USD)
2008	54000000.0
2009	239000000.0
2010	200000000.0

Table 5—Predicted data points by Gemini 1.5 Flash for Line Chart from PlotQA #21673 Portfolio Investment

2 Based on my expertise in using LLaMA 3.1 8B Instruct, we can convert among formats with almost no errors, e.g. convert prints from python code in latex table. It correctly follows instruction of to round numerical values or copy them as is.

6 REFERENCES

1. DSouza, Ashlyn, Gandhi, Kavish, and Walker, Bruce (2022). “Graph Ingestion Engine Project Report, Fall 2022”. In.