

CS 846 Software Engineering for Big Data and AI

Chat2VIS: Generating Data Visualizations via Natural Language Using ChatGPT, Codex and GPT-3 Large Language Models

Jun Lim

20870249

Various studies have been conducted to generate data visualizations from natural language texts. However, the inherent ambiguity and complexity of natural language make it difficult for traditional Natural Language Interfaces (NLIs) to accurately interpret user intent. The authors proposed a novel system (Chat2VIS), which leverages the capabilities of large language models (LLMs) like ChatGPT, Codex, and GPT-3 to convert free-form natural language into code for generating visualizations, even when queries are ambiguous or poorly specified.

The authors explored three OpenAI LLMs of the most advanced models in GPT-3, Codex and ChatGPT, which contain billions of parameters. With that, they proposed an architecture (Chat2VIS) of the following processes to convert the free-form query text into visualizations:

1. **Natural language interface:** an interface to allow users to select datasets and enter free-form texts of their intent. The visualization result is also presented in this interface.
2. **Prompt engineering:** the input dataset and intent are used to generate an LLM prompt using the “show-and-tell” technique, whose prompt is then submitted to the LLM.
3. **Script refinement and rendering:** taking the script (response) returned from the model, further refinements are done to enhance the visualization (such as labeling, coloring, etc.).

Following that, the authors conducted several case studies to evaluate the accuracy, efficiency, and robustness of the Chat2VIS in responding to various types of queries. The following findings from the six case studies were obtained:

- **RQ1:** LLM is able to support end-to-end visualization generation when supported by well-engineered prompts. These LLM could provide a few advantages, such as reducing development expenses with high response accuracy.
- **RQ2:** LLM can be effectively leveraged by engineering effective prompts using “Description and Code Prompts” or “show-and-tell” to guide and help rendering the correct chart with LLM.
- **RQ3:** The LLM’s performance among the three models does not illustrate large deviations. The only difference is that ChatGPT contains some enhanced performance, partly because they were all trained on similar datasets.
- **RQ4:** There are various limitations to using LLMs for N2LVIS, such as setting plot background color, displaying plot grid lines, specifying plot line color, variability in plot generation, and refining prompts.

Future work could include a more comprehensive evaluation of a system like Chat2VIS (such as getting user feedback) or involving benchmarking to further refine the capabilities. The authors hoped that the study could provide valuable insights for future researchers to further refine the capabilities of LLM in data visualization.

Paper Commentary

The authors provided a groundbreaking approach by leveraging LLM to generate visualizations, paving the way for future research in natural language visualizations. The paper also addressed a key challenge: Chat2VIS is capable of interpreting sophisticated and ambiguous natural language queries in visualizations, a notable achievement. Thirdly, the paper has a wide range of prompts in its evaluation, giving the tests and their results good credibility.

However, the proposed system has some downsides too. While the system is able to handle ambiguous queries, there may be limitations on processing highly complex queries that require domain-specific knowledge. Also, while the paper demonstrated its effectiveness on certain types of queries, it may not be easily generalizable across a wide range of datasets and query types (e.g., complex datasets).