Designing a database for Pokemon!

Aadya Srivastava

GenAI Tool used: ChatGPT

May 9, 2024

1 Introduction

This report documents the reflections/observations made while guiding GenAI through the process of designing a database for Pokémon, using both SQL (MySQL) and NoSQL (MongoDB) approaches.

2 Problem Statement

The task involves designing a SQL and NoSQL database for Pokémon, comprising of table creation, populating them with information provided such as Pokémon types, moves, and their attributes. The specific details include Pokémon types and their interactions, moves, and their powers. Additionally, queries need to be formulated for retrieving Pokémons capable of learning a certain move as well as moves effective against Grass-type Pokémon.

3 Prompts given to GenAI

- Objective of the first chat: To give a structured prompt in fragments so as to not overwhelm GenAI. This was done by first giving some context on what we are trying to do, then moving on to designing a database using SQL by asking it to first create the tables followed by inserting data and giving SQL commands for the two queries. A similar approach was followed to get the commands for NoSQL database.
- Objective of the second chat: To use a less structured approach to see if the GenAI's knowledge of Pokemon affects its responses.
- Objective of the third chat: To use a structured prompt but to ask GenAI to design NoSQL database first, to see if the sequence of designing databases impacts the responses.

4 Reflections on Responses Generated

- **Response to First Prompting Method:** In the initial approach, **GenAI provided accurate commands for all queries, effectively interacting with both SQL and NoSQL databases.** It demonstrated a clear understanding of database concepts and efficiently explained the differences between SQL and NoSQL technologies.
- **Response to Second Prompting Method:** In this approach, when opted for a less structured approach and given prompt to design the NoSQL database, **the initial responses were somewhat vague.** Although, they did make sense conceptually, they didn't precisely align with what I had asked for. For instance, when prompted to create collections, instead of giving an executable command, GenAI provided a general outline of how the collection structure should look like. Subsequently, when establishing the 'Pokemon' and 'Move' classic many-to-many relationship, **the method employed was not very efficient and led**

to redundant data storage, although it ultimately achieved the desired outcome. **A significant challenge encountered was GenAI's prior knowledge of the Pokémon game, which influenced its responses** until I explicitly instructed it to only utilize the provided data.

- **Response to Third Prompting Method:** In this approach, when prompted to "Create all the tables needed", in case of SQL GenAI provided an accurate query to create the necessary tables as per the prompt, however in case of NoSQL GenAI's responses tended towards offering general strategies rather than specific commands for creating collections. Despite specific prompts, NoSQL responses remained vague until directly prompted for the command to create a collection. Once this clarification was given, GenAI successfully provided NoSQL commands, utilizing the recently received input data and prompts.

  In the case of SQL, although there was some improvement compared to the previous NoSQL responses, the focus on details was still lacking. Responses tended to provide a semblance of a solution rather than a precise command and when discrepancies were pointed out, GenAI started hallucinating and adding dummy data based on its prior knowledge of Pokémon, requiring multiple attempts to generate the correct output.

  This difference was observed due to the inherent dissimilarity between SQL and NoSQL databases, particularly the absence of tables in NoSQL.

Analysis and Conclusions

Overall, GenAI does a great job when asked to stick to the provided data and when its prior knowledge of the topic doesn't interfere too much with its responses, otherwise it often starts hallucinating and adds random data to the prompts.

Also, it gives straightforward and precise responses when the prompts are clear, focusing only on one task at a time, terminology used is accurate and when context is well described.

References

- ChatGPT Transcripts
  - https://chat.openai.com/share/3c4ce08c-b58b-4cc8-a356-41968320775a
  - https://chat.openai.com/share/1e5e6a61-7e40-4c9f-90d2-bcaa73ac579a
  - https://chat.openai.com/share/abe50515-c252-4bfe-b910-170205a254b0
- Google Drive Result Link
  - SQL Query Result: https://drive.google.com/file/d/14Gec9EfcD0dwrAqQ632Y2Kw5WVfwn3-g/view?usp=sharing
  - NoSQL Query Result: https://drive.google.com/file/d/1pDE86aqxLblvO4W7hatpvGqvZZjFtf2c/view?usp=sharing