Tabletop Trove: RAG System for Board Game Recommendation

Kasey La
COSI 232 Final Presentation

01
02
03
04
05
06
07
08

INTRODUCTION
DATA
MODEL ARCHITECTURE
EXPERIMENTAL SETUP
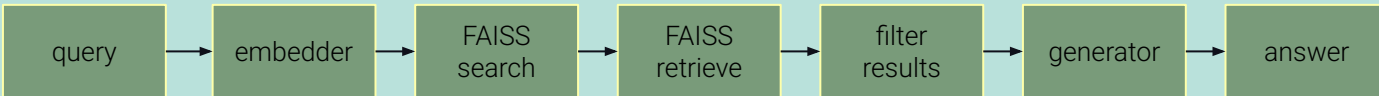RESULTS
DISCUSSION
LIVE DEMO

Introduction

- **Project goal**
  - Design a recommendation system that can help users choose a board game based on their preferences via query and filters.
- **Motivation**
  - There are so many board games out there, can a RAG system serve as a good recommender?

INTRODUCTION

DATA

MODEL ARCHITECTURE

EXPERIMENTAL SETUP

RESULTS

DISCUSSION

LIVE DEMO

Data

- **Source**
  - Kaggle dataset from 2017 (scraped from BoardGameGeek)
- **Statistics**
  - Around 94,000 games and expansions total
  - I am only using 1,000 right now: 700 games, 300 expansions
- **Cleaning steps**
  - Extracted relevant columns: name, description, category, mechanic, min/max players, min/max playtime, min age, average rating
  - Unescape escaped characters, fix commas in lists, remove html comments
  - For numeric values, add prose

Model Architecture

- **Embedding:** MiniLM-L6-v2

- **Retriever:** FAISS

- **Generator:** Llama-3.2-1B Instruct

query → embedder → FAISS search → FAISS retrieve → filter results → generator → answer

01
02
03
04
05
06
07
08

INTRODUCTION

DATA

MODEL ARCHITECTURE

EXPERIMENTAL SETUP

RESULTS

DISCUSSION

LIVE DEMO

Experimental Setup

LLM prompting

- ○ At first, the retrieved passages were given in list format: the game's details were all listed one after the other separated by newlines (think JSON).
  - ■ **Issue:** LLM sometimes tried to continue the pattern and created a new game instead of summarizing one in the retrieved passages.
- ○ To combat the problem, the game's details were rewritten to be in prose/paragraph format instead.

Results

- **Passages in list format**
  - Most of the time, the generated answer was complete garbage, either repeating the prompt or creating a new game in the list format.
- **Passages in prose/paragraph format**
  - Generated answer improved significantly:
    - No more repeating the prompt
    - No more coming up with new games
- **Both formats**
  - If no hallucinations, the answers looked decent.

Discussion

- **Issues**
  - Retrieval does not always get the games that I am expecting
- **Future directions**
  - Do proper evaluation with ROUGE and/or human evaluation
  - Utilize more of the data
  - Use a larger model to see if performance improves

Live Demo