Building an Intelligent Conversational Agent

Running Hou 22735591

Abstract

Abstract In this project, we implemented a conversational agent which answers natural language questions and interacts in a conversational way with humans. The dialogue format makes it possible for the agent to answer different types of questions, including factual questions, embedding questions, recommendations, and multimedia inquiries. We also introduced crowd-sourcing to improve the data quality.

Keywords: *Knowledge Graph; Embedding; Multimedia; Recommendation; Crowd-sourcing*

1. Methods

To answer the factual questions, we used a pre-trained NER transformer model to recognize the movie names. Then we labeled the movie names as ENTITY and got the relationship in a question. We matched the triple, which contains a subject, a predicate, and an object. Finally, we extracted information from the Knowledge Graph of Wiki data using the Resource Description Framework(RDF) query language.

We also provided another way to answer the factual questions by calculating the similarity distances of entities and relations using the embedding data. For a recognized entity and its relation, our first try was to find the 2nd entity through triple matching. However, in questions like 'Who is the screenwriter of The Masked Gang: Cyprus?', the movie does not have a screenwriter relation. Under this situation, we managed to the most similar relation and got the corresponding 2nd entity. Then we calculated the similarity distances list of the 2nd entity and returned the top 3 results.

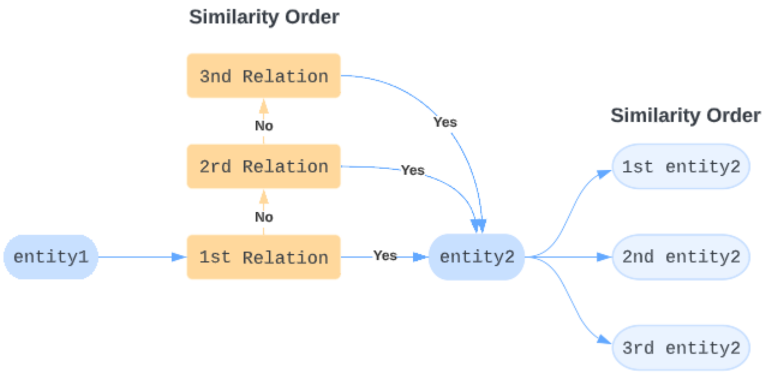


Figure 1: Embedding answers design

We designed a common relations filtering mechanism to generate a recommendation list with a group of given movies. Its main idea was to find the common (relation)-(2nd entity) pairs among the movies provided in the question. For each pair, we turned back and managed to find all the 1st entities. Then we gather them in a list and pick the top 3 in frequency order. The intuition is to extract the common characteristic list; then, we can find the movies which own these characteristics, and the more, the better.

We gave the agent access to multimedia data and crowd-sourcing data to help it provide better responses. For the multimedia data, we first got an entity's IMDB id(P345). Then we returned the figure code by searching the multimedia JSON file. To improve the answer quality and avoid the confusion caused by multi-object figures, we returned the figures with only one cast for the questions asking to show a specific cast. To augment and correct the data in the knowledge graph, we processed the crowd-sourced data and merged it with the initial knowledge graph. To be specific, we deleted data that had nonsense response time and accuracy from malicious crowd workers and aggregated individual crowd answers with majority voting. Finally, we computed the inter-rater agreement (Fleiss Kappa) per batch.

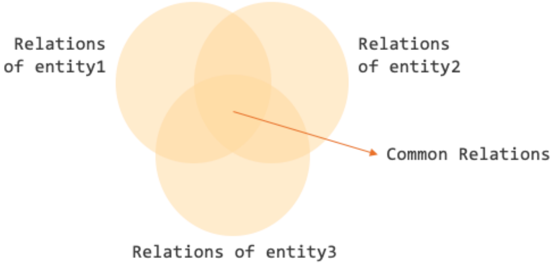


Figure 2: Common characteristics of movies

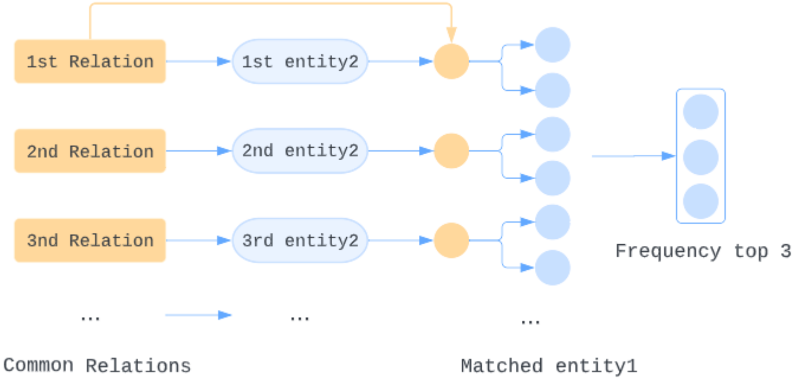
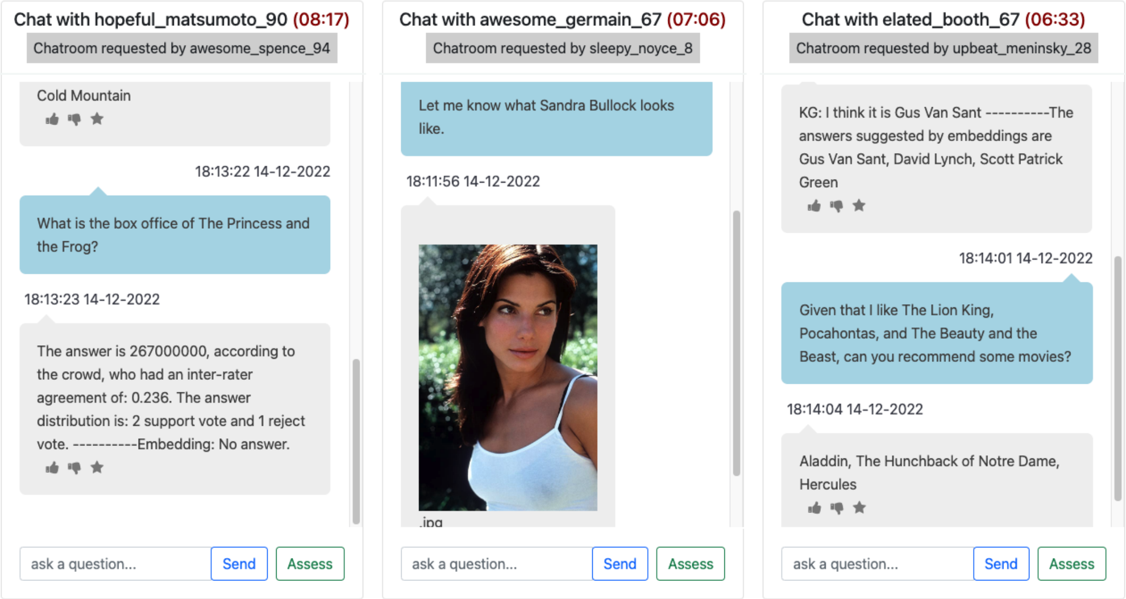


Figure 3: Recommendation answers design

To create a multi-functional agent for intelligent conversation, we needed to collect human-agent interactions, which consisted of five types of questions to test the model response quality. To collect these interactions, we took conversations that AI trainers had with the chatbot. We randomly selected a piece of information from the knowledge graph, encapsulated the elements into a question, and had AI trainers ask it. Using these questions, we can improve the model's generalization ability and robustness. We evaluated the agent using a web-based infrastructure, Speakeasy, via HTTP requests.

2. Results

In the following samples, the chat agent is asked in different ways. We merged the crowd-sourced data with the KG answers, so if a piece of information is found in the crowd data, the KG method will output a crowd-sourced result with votes and the inner rate. For other factual questions, the agent will output answers from both KG and embeddings to show the technological difference. For the recommendation questions, the agent will answer 3 movie names. And for the multimedia question, it will post a picture. If the agent fails to find an answer, it will simply say 'No answer', as shown in figure4 on the next page.



Chat with hopeful_matsumoto_90 (08:17)

Chatroom requested by awesome_spence_94

Cold Mountain

18:13:22 14-12-2022

What is the box office of The Princess and the Frog?

18:13:23 14-12-2022

The answer is 267000000, according to the crowd, who had an inter-rater agreement of: 0.236. The answer distribution is: 2 support vote and 1 reject vote. ----------Embedding: No answer.

Chat with awesome_germain_67 (07:06)

Chatroom requested by sleepy_noyce_8

Let me know what Sandra Bullock looks like.

18:11:56 14-12-2022

.jpg

Chat with elated_booth_67 (06:33)

Chatroom requested by upbeat_meninsky_28

KG: I think it is Gus Van Sant ----------The answers suggested by embeddings are Gus Van Sant, David Lynch, Scott Patrick Green

18:14:01 14-12-2022

Given that I like The Lion King, Pocahontas, and The Beauty and the Beast, can you recommend some movies?

18:14:04 14-12-2022

Aladdin, The Hunchback of Notre Dame, Hercules

Figure 4: Question samples on Speakeasy