Research on Latent Semantic Relationship Search Engine Based on Knowledge Graph

Minqin Mao(✉) and Jielan Zhang

Jiangxi Teachers College, Yingtan 335000, Jiangxi, China
maominqin@126.com

Abstract. Knowledge graph is a large database composed of entities, relationships and attributes, which can provide rich semantic information for search engines. The potential semantic relation search engine based on Knowledge graph is a novel search engine. It obtains potential semantic relationships from the Knowledge graph, and then uses these potential semantic relationships to search for data sources such as web pages and documents. This paper first analyzes the characteristics of the Knowledge graph, then lists the construction process of the Knowledge graph based on WordNet, and finally proposes the potential semantic relationship search engine architecture based on the Knowledge graph.

Keywords: Knowledge graph · Potential semantic relationships · Search Engines

1 Introduction

At present, search engines lack in-depth semantic analysis and understanding of massive text data. In the search process, it is often only possible to obtain corresponding data from the text, rather than extracting potential semantic relationships from it. Using this method of retrieval, it is not only difficult to meet user needs, but also leads to inaccurate query results.

The related search technology based on the potential semantic relationship of the Knowledge graph emerged in this environment. This technology extracts potential semantic relationships from the Knowledge graph to facilitate semantic related retrieval by users.

2 Definition and Characteristics of Knowledge Graph

The Knowledge graph is a knowledge base that forms a structured model based on logical abstraction of massive amounts of information from the real world. It is a large database consisting of entities, relationships, and attributes, which can describe an entity and its associated information. Knowledge graph is essentially a semantic network. The nodes represent entities, the vertices represent attributes, and the edges represent various semantic relationships between the entity attributes. The process of constructing Knowledge graph involves three stages: entity extraction, attribute extraction, and relationship extraction.

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2024
Published by Springer Nature Switzerland AG 2024. All Rights Reserved
C. Wu et al. (Eds.): MONAMI 2023, LNICST 559, pp. 269–274, 2024.
https://doi.org/10.1007/978-3-031-55471-1_20

The Knowledge graph has the following characteristics.

1. The Knowledge graph uses the form of graphs to represent various elements and the relationships between them in the real world. It is an abstract expression of the real world.
2. Knowledge graph is a knowledge base that is made up of entities, relations, attributes, and more. It is based on facts and describes the objective world in a dynamic way and with a large amount of information. It is highly scalable and can transform human knowledge and experience into a form that can be understood by computers, making it highly versatile.
3. There are a lot of potential semantic relations in Knowledge graph, and different entities may have some connections. The connection may be subjective or objective. For example, the relationships among entities include hierarchical relationships, sibling relationships, adjacency relationships, inclusion relationships, and so on. These potential semantic relationships can be modeled to help us better understand the relationships between data and make more effective and accurate decisions.

3 Construction Process of Knowledge Graph

Building a Knowledge graph with WordNet dictionary can reduce the number of steps required to structure data. The construction process based on WordNet Knowledge graph includes the following steps:

1. **Clarify the purpose and scope of the Knowledge graph**: Before building a Knowledge graph, you need to clarify its purpose and scope, so as to better organize and manage information.
2. **Data collection:** Take the data source in WordNet as the original data, and determine the core concepts of sub domains by manually dividing the identified knowledge domains.
3. **Knowledge Extraction:** Extract knowledge entities, relationships, and attributes related to core domain concepts from structured, semi-structured, and unstructured data of WordNet. The acquired data are integrated to form a preliminary knowledge representation.
4. **Building Relational Networks:** Using the synonymous, antonymous, hypernymy, and partial overall relationship in WordNet2. Knowledge entities and relationships are gradually built into an entity relationship network through deep learning algorithms and natural language and other technologies.
5. **Storage and Management:** The extracted information is stored in the database, and the information in the database is regularly updated to ensure its accuracy, integrity and reliability.
6. **Update and maintenance:** We need to update and maintain the Knowledge graph regularly, and integrate the extracted structured data and new data to generate a more complete and accurate Knowledge graph.

4 The Architecture of Latent Semantic Search Engine Based on Knowledge Graph

Potential semantic search module traditional search engines retrieve web pages or documents by keyword matching. For example, search engines match search websites based on keywords and return web pages related to the search keywords. However, for many different web pages, traditional search engines will return multiple results. Traditional search engines are just a simple resource retrieval tool, which cannot effectively discover potential semantic relationships in data resources. In order to solve these problems, the author proposes a potential semantic relationship search engine architecture based on Knowledge graph. The search engine is a new type of search engine based on Knowledge graph. It extracts valuable information from a large amount of data and transforms it into a form that users can understand. The basic process of this framework system is as follows (Fig. 1).

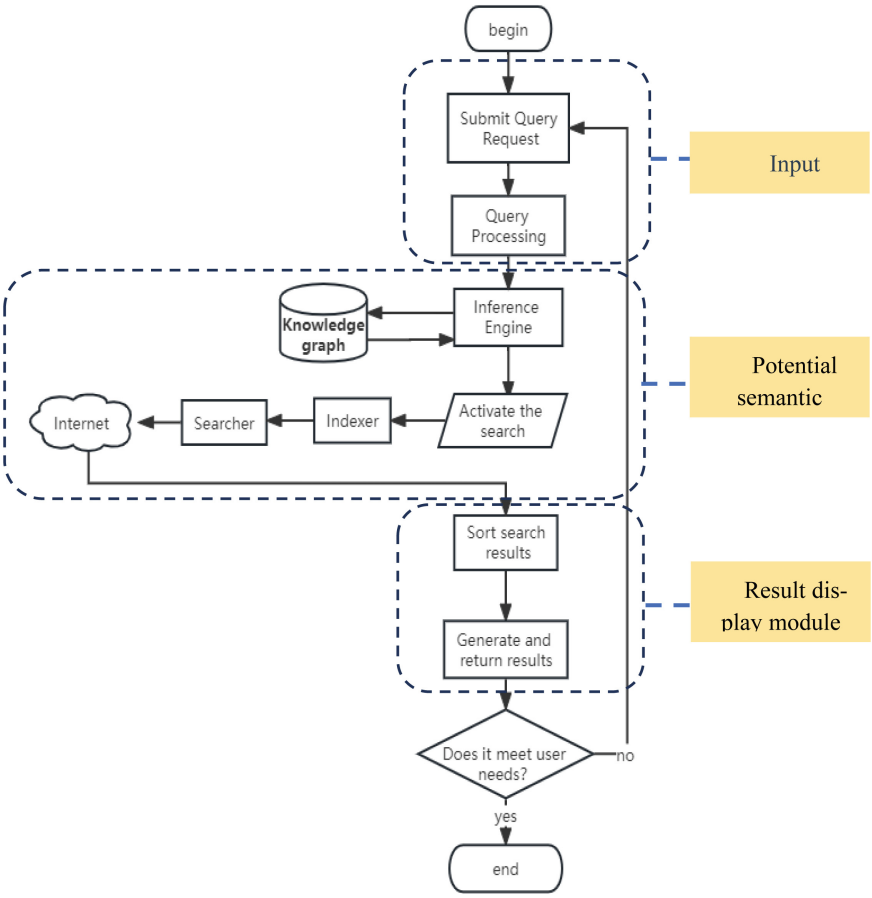

Fig. 1. Architecture diagram of a new search engine system based on knowledge graph

1. **Input module:** The user submits the query request through the query interface of the search engine, performs query processing, and submits the results of query processing to the inference engine.
2. **Knowledge graph building module:** This module uses natural language processing technology and machine learning algorithm to build a Knowledge graph, uses existing domain knowledge, domain model, etc. Through training and optimization of features, converts them into meaningful semantics, and realizes effective understanding and application of knowledge. It can contain a large amount of entity and relationship data, such as entities, attributes, relationships between attributes, etc. Then these data and relationships are transformed into structured representation, stored in the database, and the information in the database is regularly updated to ensure its accuracy, integrity and reliability.
3. **Potential semantic search module:** This module uses inference rules, statistical methods or machine learning algorithms in inference engine to extract entities and relationships from Knowledge graph, and classify and interpret them. The inference results from the inference engine are sent to the searcher, which provides high quality search results through various retrieval techniques and algorithms, such as information retrieval and rule-based query optimization, to provide high-quality retrieval results.
4. **Result display module:** The module is implemented using graphical a user interface (GUI) or a natural language generation tool (NLG). The search results and related information, such as text, image, audio and other multi-modal information, are integrated into a visual form that can be browsed and queried. Index and sort algorithms are used to sort the retrieved results in order. The sorted results are submitted to the page generation module to generate a result page, and the user accesses the generated search results through the access interface to improve the search effect and user experience.

5 Summary

The potential semantic relation search engine based on knowledge graph can provide rich semantic information for users. Its main advantage lies in the integration of Knowledge graph and potential semantic relation. Although the potential semantic relationship search engine based on Knowledge graph has made some achievements, it still needs to be further improved.

1. For example, the existing Knowledge graph cannot meet the requirements of large-scale data retrieval. How to add more entity, relationship a attribute information to the Knowledge graph is also a problem worth studying.
2. Although the potential semantic relation search engine based on Knowledge graph can use the information in knowledge atlas to some extent, it still needs manual annotation to get high quality potential semantic relation. There are three reasons. First, the Knowledge graph contains a lot of noise information. Second, some knowledge in the Knowledge graph is not completely accurate. Third, some data sources may not exist in the knowledge base.

3. How to use existing technology to improve search performance is also a challenge. There is currently no unified and clear answer on how to use existing technologies to improve the performance of latent semantic relationship search engines, and further research and exploration are needed.

References

1. Investigators at Nanyang Technological University Detail Findings in Neural Networks and Learning Systems (Brain-inspired Search Engine Assistant Based On Knowledge Graph). 2022(Jan.17), 13–14
2. Greg, R.: Notess. Search engine to knowledge engine? Online Search. **37**(4), 61–63 (2013)
3. Mayank, K., Pedro, S.: Knowledge graph for social good: an entity-centric search engine for the human trafficking domain. IEEE Trans. Big Data **8**(3), 592–606 (2022). https://doi.org/10.1109/TBDATA.2017.2763164
4. Uyar, A., Aliyu, F.M.: Evaluating search features of Google Knowledge Graph and Bing Satori Entity types, list searches and query interfaces. Online Inf. Rev. **39**(2), 197–213 (2015). https://doi.org/10.1108/OIR-10-2014-0257
5. Asgari-Bidhendi, M., Hadian, A., Minaei-Bidgoli, B.: FarsBase: the Persian knowledge graph. Semant. Web **10**(6), 1169–1196 (2019). https://doi.org/10.3233/SW-190369
6. Du, Y., Li, C., Hu, Q., et al.: Ranking webpages using a path trust Knowledge graph. Neurocomputing **269**(Dec.20), 58–72 (2017). https://doi.org/10.1016/j.neucom.2016.08.142
7. MICROSOFT TECHNOLOGY LICENSING, LLC. Knowledge graph for conversational semantic search:US15664124. 2022-09-06
8. Ma, C., Zhang, B.: A New query recommendation method supporting exploratory search based on search goal shift graphs. IEEE Trans. Knowl. Data Eng. **30**(11), 2024–2036 (2018). https://doi.org/10.1109/TKDE.2018.2815544
9. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American (2001)
10. Gruber, T.H.: A translation approach to portable ontology specifications. Knowl. Acquis. **2**, 199–220 (1993)
11. Han, L., Finin, T., Joshi, A.: GoRelations: an intuitive query system for DBpedia. Semant. Web **04**, 674–693 (2012)
12. Li, Q., Yang, W., Ye, X., Ma, X.: Research on knowledge base of device test training system based on rough set data mining. In: Proceedings of the 2013 International Conference on Intelligent System, Applied Materials and Control Technology (GSAMCT 2013) (2013)
13. Bordes, A, Usunier, N., Garcia-Duran, A., Weston, J., Yakhnenko, O.: Translating embeddings for modeling multi-relational data. In: Proceedings of the NIPS (2015)
14. Momtchev, V., Peychev, D., Primov, T.: Expanding the pathway and interaction knowledge in linked life data. In: Proceedings of International Semantic Web Challenge (2015)
15. Meng, Z.: Research on Construction of Course Knowledge Graph and Search Technology. Dissertation for Doctor Degree of Wuhan University (2016)
16. Xiangqian, L.: A method of searching entities based on wordnet noun network. Dissertation for Master Degree of Nanjing University (2015)
17. Yuncheng, G.: Research on Chinese-English-Mongolian Term Knowledge Graph of Computer Field Based on WordNet. Dissertation for Master Degree of Inner Mongolia Normal University (2021)
18. Berthold, M.R.: Towards bisociative knowledge discovery. Bisociative Knowledge Discovery. Springer-Verlag, Berlin, Heidelberg (2012)

19. Tom, H., Christian, B.: Linked Data. Morgan & Claypool., San Rafael (2011)
20. Berasaluce, S., Laurenço, C., Napoli, A., Niel, G.: An experiment on knowledge discovery in chemical databases. In: Boulicaut, J.F., Esposito, F., Giannotti, F., Pedreschi, D. (eds.) Knowledge Discovery in Databases: PKDD 2004. PKDD 2004. LNCS, vol. 3202. Springer, Berlin, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30116-5_7