COMPARING AI ALGORITHMS FOR CUSTOMER SUPPORT SYSTEMS BASED ON EXISTENT KNOWLEDGE BASE

Ivan Kiprich, student,
M. Suknov, PhD, Associate Professor,
Kharkiv National University of Radio Electronics

Customer support systems play a vital role in modern businesses, ensuring efficient communication and problem resolution between customers and companies. Existing Knowledge base in customer support systems comprises a repository of information that contains frequently asked questions, troubleshooting guides, and other relevant documentation. The knowledge base serves as a valuable resource for customer support agents and AI algorithms in retrieving appropriate information to address customer queries effectively. With the advent of artificial intelligence (AI), OpenAI algorithms have emerged as powerful tools to enhance customer support experiences. This article aims to compare various OpenAI algorithms for customer support systems, specifically focusing on their ability to leverage an existing knowledge base.

Traditional ways to organize a knowledge base include “Hierarchical”[3], “Sequential”, “Tag-based”, “Graph”, “Faceted” and “Hybrid” structures. Each structure has its own limitations and its usage requires a specific approach. Common problems associated with searching data in hierarchical knowledge bases are: over-reliance on navigation, lack of contextual answers, limited flexibility and user familiarity and training. For tag-based knowledge bases there are: tag inconsistency, tag overload, lack of hierarchy and context and limited query precision.

With the evolution of open data, better information extraction frameworks and crowdsourcing tools, large-scale structured knowledge bases are becoming more available. This data can be used to provide common sense knowledge for semantic applications. However, querying and reasoning over this data demands approaches which are able to cope with large-scale, semantically heterogeneous and incomplete KBs [1].

There are a number of AI algorithms used for search in a knowledge base.

Commonly utilized Vector Space Model (VSM), Latent Semantic Indexing (LSI)

The idea of the VSM is to represent each document in a collection as a point in a space (a vector in a vector space). Points that are close together in this space are semantically similar and points that are far apart are semantically distant [2]. VSM relies on the term frequency-inverse document frequency scheme to assign weights to terms in documents and queries. The cosine similarity measure is then used to rank the similarity between vectors. VSM has several attractive properties: simplicity, efficiency, language independence, and document ranking. VSM's simplicity and ease of implementation make it a popular choice for search algorithms in customer support systems.

Latent Semantic Indexing (LSI) tries to overcome the problems of lexical matching by using statistically derived conceptual indices instead of individual words for retrieval [4]. LSI aims to capture the hidden relationships between terms and documents by applying dimensionality reduction techniques, such as singular value decomposition (SVD). LSI overcomes the limitations of the exact term matching in search queries and enables the retrieval of documents based on their semantic similarities rather than precise keyword matches.

VSM performs well when the focus is on exact term matches, making it suitable for scenarios where customers expect precise responses. On the other hand, LSI's ability to capture semantic relationships makes it better equipped for handling ambiguous queries or retrieving relevant documents based on contextual similarity. LSI excels in scenarios where customers may not be aware of the exact terminology used in the knowledge base. LSI outperforms VSM in scenarios where the queries are contextually complex and when the knowledge base contains a significant amount of domain-specific jargon.

While VSM's simplicity and efficiency make it a reliable choice for precise term matching, LSI outperforms VSM in capturing semantic relationships and handling complex queries. The choice between VSM and LSI depends on the specific domain, knowledge base amount and query complicity.

References

1. A Distributional Semantics Approach for Selective Reasoning on Commonsense Graph Knowledge Bases : Natural Language Processing and Information Systems, 2014. - 21-23
2. From Frequency to Meaning: Vector Space Models of Semantics : Journal of Artificial Intelligence Research, 2010. - 141-188
3. Hierarchical Knowledge Bases and Efficient Disjunctive Reasoning : Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning/Ronald J. Brachman, Hector J. Levesque, Ray Reiter M. Kaufmann, 1989. - 33-43
4. Latent Semantic Indexing: An overview INFOSYS 240 Spring 2000

INVESTIGATION OF THE EFFICIENCY DEPENDENCE OF RELATIONAL AND GRAPH DATABASES ON DATA

Nedaiev I. S., student,
Gubaryeva O. S., PhD, Associate Professor,
Kharkiv National University of Radio Electronics

There are various types of databases, each with its own purpose, strengths and weaknesses. Since each type specializes in solving its own specific task, it is impossible to determine which one is the best overall. However, if a specific situation is considered, it is possible to determine which type of database is better suited for that particular case.

In general, many DBMSs support different data types, and almost any information can be converted to the appropriate format one way or another. However, such solutions can greatly affect the speed and quality of their work, or even their security and fault tolerance. For example, relational databases are considered a classic and most information can be represented as tables and relationships between them [7]. However, complex relationships significantly affect performance and resource consumption, as well as complicate the architecture of the database itself, which can lead to its malfunction. On the other hand, other types of databases may be better suited for such situations but will not be able to reproduce a simple scenario of using