Scholar Publishing Group
Frontiers in Educational Psychology
https://doi.org/10.38007/JEP.2024.050108
ISSN 2790-0975 Vol. 5, Issue 1: 62-70

Educational Psychology Question Answering System Based on BERT Pre-training Language Model

Kang Wang[1,a,*]

[1]*The College of Information Engineering, Wuchang Institute of Technology, Wuhan, Hubei, China*
[a]*156942090@qq.com*
[*]*Corresponding author*

Keywords: Question-Answering System; Knowledge Graph; BERT; BiLSTM-CRF

Abstract: In this paper, we investigate an intelligent question-answering system based on knowledge graphs. We first discuss the construction of knowledge graphs, the entity recognition scheme, and a novel question parsing algorithm models. Specifically, we employ the BiLSTM-CRF model for named entity recognition, which integrates Bidirectional Long Short-Term Memory Networks (BiLSTM) and Conditional Random Fields (CRF) to significantly enhance recognition accuracy and efficiently capture the contextual information. For question parsing, We use the BERT pre-trained language model, which can achieve an exact matching between questions and templates based on the high-dimensional vector encoding. In order to enhance the system performance and user experience, we introduce the Model-View-Controller (MVC) scheme to optimize our proposed system architecture. In particular, the data layer is responsible for data access, the logic processing layer handles algorithm and logic processing, and the user interaction layer provides a friendly interface and interaction method. The experiment results show the proposed intelligent question-answering system has made significant progress in entity recognition and question parsing and can achieve efficient operation and user-friendly interaction experience.

1. Introduction

With the rapid development of artificial intelligence (AI) technologies, many natural language processing (NLP)[1] models have proposed for text understanding and information extraction. Pre-trained language models have become the powerful tools for solving complex language understanding tasks, where the pre-trained language model can capture the deep linguistic patterns and semantic information based on the data pre-training[2]. These models has greatly pushed the boundaries of language processing technology, making deep understanding and application of natural language possible.

Educational psychology[3], as a discipline that studies human psychological activities, psychological characteristics, and behavioral laws in the process of education, occupies an important position in the field of education. With the rise of online education, a large amount of education-related text data is generated. Thus, how to effectively extract and parse psychological interrogative sentences is of great significance for improving the quality of education and assisting teaching decisions. However, the complexity and specialization of educational psychology texts bring severe challenges for parsing the interrogative sentences. Traditional rule-based and simplistic machine learning methods[4] face the limitations due to their inadequate comprehension of the underlying semantics of language when dealing with these challenging problems. In addition, educational psychology texts contain a large number of specialized terms and complex sentences, which make it difficult for traditional methods to accurately parse interrogative sentences.

In this context, we propose an educational psychology question and answer system based on the BERT pre-trained language model. Focusing on the design of a Q&A system in an educational psychology course, our work aims to deeply analyze the linguistic features of educational psychology texts by utilizing the powerful language comprehension capability of the BERT model. The experiment results show the accuracy and stability of the system on educational psychology Q&A. Our proposed system aims to provide an efficient and personalized learning experience. Thus, this work provides innovative perspectives on how to teach and learn in educational psychology courses, and provides practical examples and theoretical support for the application of intelligent educational technologies.

2. Related Work

Recently, many novel Q&A systems have been proposed[5]. Benefiting from advanced technologies, such as NLP, machine learning, and deep learning, the comprehension and response capabilities of Q&A systems have been significantly improved. For example, the utilization of deep learning models for semantic comprehension of questions can more precisely capture user intention and contextual information. In addition, for efficient querying and reasoning of knowledge graphs[7], researchers have proposed various optimization algorithms and frameworks to improve the accuracy and responsiveness of Q&A systems. Also, many researches in the field of knowledge graph-based Q&A systems have also made remarkable progress, covering a wide range of aspects such as the construction of knowledge graphs[8] and the design of Q&A systems. There exist some literatures regarding the researches on the Chinese based knowledge graphs such as CN-DBpedia[9] of Tsinghua University, the Knowledge Workshop[10] of Fudan University. These pervious works provide mature technical support for Chinese Q&A systems. For the challenging problem of structuring accuracy knowledge graphs, many researchers have also proposed a variety of effective solutions, such as The DL based automatic complementation technique[11] and error correction of knowledge graphs[12]. In addition, the researches on the implementation of Q&A systems also have achieved important result[13]. Many pervious works focus on the advancement of NLP technologies, such as semantic understanding and contextual relationship recognition[14], to enhance the comprehension and response capabilities of Q&A systems. For the challenge of large-scale knowledge graphs, researchers have optimized query algorithms, thus improving the retrieval efficiency and accuracy of Q&A systems. For example, some Q&A systems can achieve high accuracy for the complex questions by introducing graph neural networks[15].

3. Intelligent Q&A System Design

3.1. General framework

In this paper, we use the educational psychology knowledge graph as a knowledge base and propose a novel model for the Q&A system that aims to acquire relevant knowledge of the input questions. The framework of the proposed model contains the NLP techniques, the knowledge graph construction, and the query methods. Then, the users can obtain the educational psychology related knowledge and the answers by entering the corresponding questions to the proposed model.

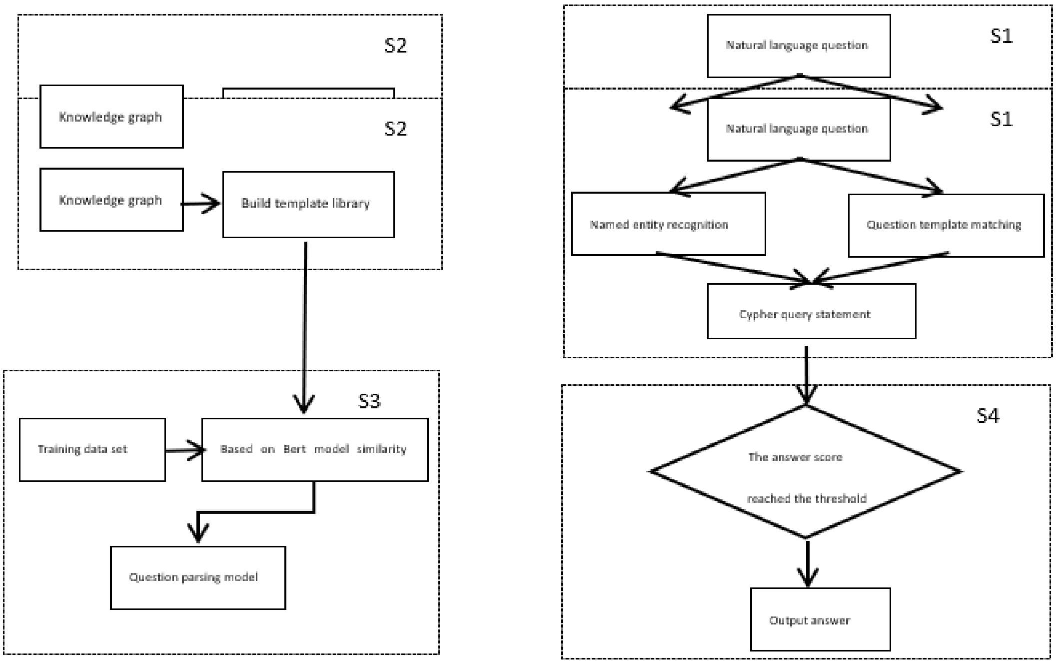

Figure 1: Flowchart of the Q&A system.

The proposed Q&A system based on educational psychology knowledge contains five modules:

(1) Entity Recognition Module: The named entities in the question and answer are recognized and extracted using the named entity recognition model based on BiLSTM-CRF. (S1)

(2) Template construction module: By utilizing knowledge graphs and educational psychology related questions, combined with regular expression syntax, we are able to construct generic templates for questions that cover different types of questions and capture the key information in the questions. (S2)

(3) Question Parsing Modeling Module: Build a BERT-based text similarity model, process the pre-organized collection of Chinese similarity network public data for model training. By using the pre-trained BERT language model , our proposed model can encode the questions and templates of user input, thus extracting their semantic information. It can map the questions and templates to the high-dimensional vectors, making them amenable to comparison and similarity computation, and realizing the parsing of interrogative sentences into Cypher query statements. (S3)

(4) Answer selection module: If the highest confidence level of the list of answers obtained from the retrieval in the knowledge graph exceeds the set threshold, the system will directly return that answer as the final result, such a strategy ensures accuracy and reliability. (S4)

3.2. Entity Recognition Module

The named entity recognition model based on BiLSTM-CRF mainly contains 5 layers such as data preprocessing layer, word vector conversion layer, BiLSTM layer, CRF layer and model output.

In terms of specific operation, firstly, the data pre-processing layer segmented the text and removed the deactivated words to prepare for feature extraction. Second, the word vector conversion layer converts the text to word vectors using pre-trained word vectors such as Word2Vec to capture semantic and contextual information. The BiLSTM layer learns the feature information of the input sequences through forward and backward LSTM and synthesizes the feature vectors. The CRF layer establishes the global dependency between labels through the transfer matrix and performs accurate label prediction. Finally, the output layer decodes the CRF output to obtain the final named entity recognition result. This model provides efficient and accurate named entity recognition for natural language processing, which has important application value.

3.3. Template Construction

Templates are crucial in natural language question parsing, especially in mapping predicates to relations or attributes of the knowledge graph, which has a direct impact on the performance of the Q&A system. Based on the knowledge graph in the field of educational psychology, manual and automated methods are used to generate one-to-many templates by combining regular expression grammars that contain regular expressions, related attributes, and prioritized fields to accurately match the relations and attributes in the questions. Generating templates involves organizing the relationships and attributes in the knowledge graph and analyzing a large number of questions with a syntactic analysis tool to identify keywords and transform them into matchable template forms. This process not only improves question parsing capabilities, but also effectively supports the operation of the Q&A system, enhancing its efficiency and accuracy.

The usage of each field is described below:

(1)Content field: describes the content of the template and is written using regular expression syntax. In the template, named capture groups are used (? <Method>(*)?) which is a form of determining the subject. For example, in the question "What is the research methodology of educational psychology?" When matched to this template, the captured phrase is "of educational psychology". Removing the stop word gives the subject "educational psychology".

(2)Subject field: identifies whether the subject of the question is certain or not, defaults to true. when the subject is uncertain, it is identified as false. e.g. "Who is known as Erdu?" is a case where the subject is uncertain. The purpose of setting this field is to subsequently determine the syntactic structure of the query statement.

(3)Value field: indicates whether the object is certain.

(4)Type field: indicates the relationship or attribute corresponding to this template. In KM, relationship refers to the "edge" that connects two entities, e.g., the relationship between "China" and "Beijing" is "capital". ". Attributes are the knowledge of the entity itself, for example, one of the attributes of "China" is "area", and the corresponding value of the attribute is "16,800,000 square kilometers".

(5)Class field: indicates the type of the subject of the question. It is used to qualify the type of subject in some special questions. class field includes types such as time, person, etc., which are mostly empty and are mainly used to identify the type of subject in a specific field.

(6)Usage field: used for special handling identification. Since some types of questions cannot get results directly through Cypher query, such questions need to be handled specially.

(7)Priority field: indicates the priority of the template, which is mainly used to calculate the score of the predicate. There are three priorities: the first priority is the template generated based on the predicates of the question, the relationships or attributes in the knowledge graph, and the specific type of question targeted with a high confidence level. The second priority is templates generated using question words with distinctive features, mainly used to match some questions about attributes when the first priority templates cannot be matched. The third priority is the use of broader question words to match when neither the first nor the second priority can be matched.

3.4. Template Matching

In the proposed Q&A system, the key issue is how to match the questions asked by users with the constructed question templates. In this paper, we adopt the cos-similarity computation method, which first encodes the questions and templates as vectors using the pre-trained BERT language model, represents them as high-dimensional vectors, and then evaluates the similarity between them by the cosine angle computation method. Cos-similarity measures the degree of similarity between two vectors by calculating the cosine of the angle between them. Finally, the questioning templates are sorted according to the similarity score and the template with the highest similarity to the user's question is selected for matching. This approach improves the accuracy and efficiency of the system and provides effective question parsing and answering support for Q&A systems.

Using the cosine similarity calculation method, it can accurately measure the semantic similarity between the question and the template, and provide effective question parsing and answer support for the Q&A system. With this matching method, the accuracy and efficiency of the system can be improved to further optimize the performance of the Q&A system.

$$S_{\text{cos-similarity}} = \frac{V_{c_i} \times V_{c_j}}{||V_{c_i}|| \times ||V_{c_j}||}$$

where V_{C_i} and V_{C_j} are vectorized representations of string C_i and string C_j, respectively; $||V_{C_i}||$ and $||V_{C_j}||$ denote the vector lengths of string C_i and string C_j, respectively.

4. Data Selection and System Simulation

4.1. Data Selection

(1) BERT pre-training

The training dataset is fed into the BERT model to encode Chinese interrogative sentences and extract deep semantic features. The test set and training set samples contain a large amount of text data, which may come from multiple sources such as the Internet, books, news, forums, etc., to ensure that the model has the ability to understand language in multiple domains and topics.The distribution states of the test dataset for the BERT model are as follows.

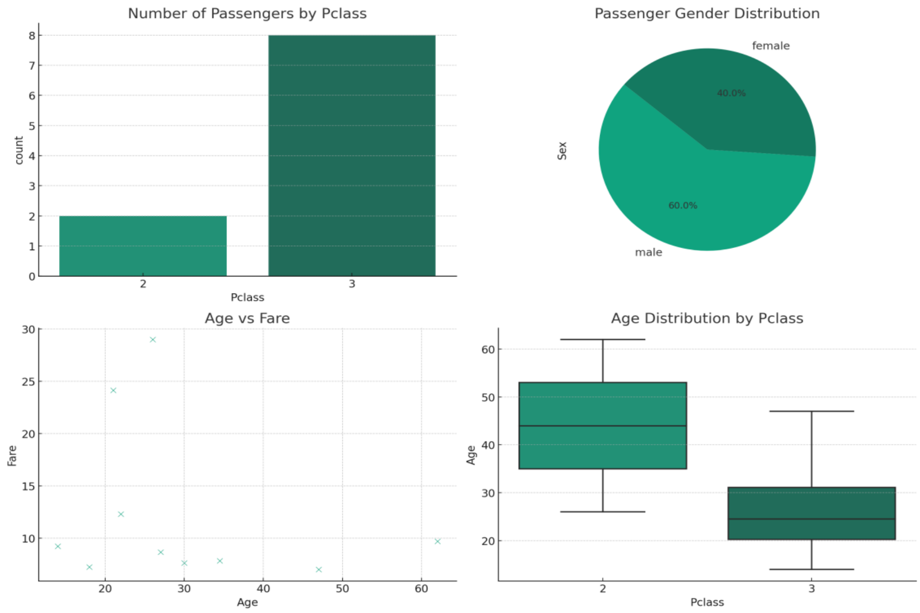

Figure 2: Data distribution of Bert's model test set.

(2) System data

Text data annotation, the key information in the text is labeled through the annotation work in the professional field to prepare for the subsequent knowledge extraction. Features such as educational psychology knowledge points and entity relationships are extracted from the text through knowledge extraction methods. The extracted knowledge entities and relationships are stored through storage tools (e.g., Neo4j graph database) to form a complete knowledge graph. Users can retrieve entities directly through the search interface, and the graph mapping shows the relationships of educational psychology knowledge. In the relationship mapping, boxes represent concepts, circles represent entities, arrows represent relationships, and circles of the same color indicate that the instances correspond to the same concepts. The most important feature of this knowledge mapping is that it shows the relevant information about the retrieved knowledge point and provides the location of the knowledge point in the educational psychology textbook, which is convenient for teachers and students.

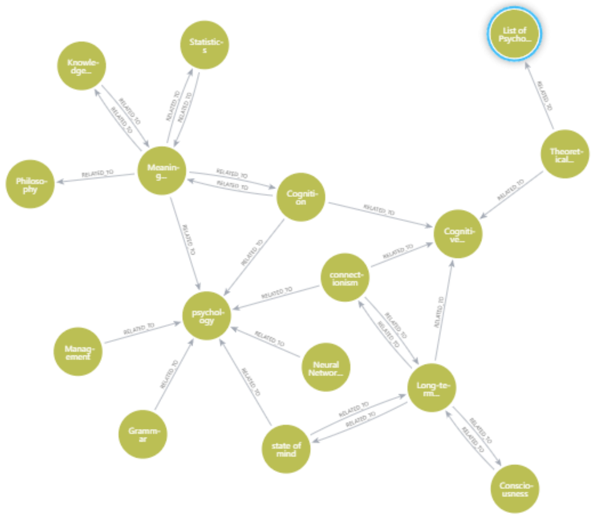

Figure 3: Educational Psychology Knowledge Mapping.

4.2. Template Matching

How to match between the questions and answers asked by the user is the core of the Q&A system, and in addition to using the similarity scores mentioned above, there are other quantitative metrics such as accuracy, precision, recall, and F1 scores.

When evaluating the effectiveness of template matching in a question and answer system, accuracy indicates the proportion of templates correctly selected by the matching system, i.e., the number of correctly matched samples as a proportion of the total number of samples. Accuracy reflects the overall matching accuracy of the system, for example, for the question "What are the research methods in educational psychology?" For example, for the question "What are the research methods in educational psychology?", the corresponding expected answer is "The research methods in educational psychology include field research, experimental research, and literature analysis." If the actual response is "Research methods in educational psychology include fieldwork, experimental research, and literature analysis." then the accuracy assessment would be "The expected response is consistent and accurate with the actual response." Precision and recall are metrics used to evaluate matching performance in more detail. Precision indicates the proportion of the number of positive samples correctly matched to the number of all matched samples, with the formula:

$$\text{Precision} = True\ Positives/(True\ Positives + Fales\ Positives)$$

Where True Positives indicates the number of correctly matched samples and False Positives indicates the number of incorrectly matched samples.

Recall represents the ratio of the number of correctly matched positive samples to the number of all actual positive samples, and is given by the formula:

$$\text{Recall} = True\ Positives/(True\ Positives + False\ Negatives)$$

where False Negatives indicates the actual number of positive samples that failed to match. These two metrics combined provide a more complete picture of the performance of the matching system.

The F_1 score is a combination of precision and recall metrics that evaluates matching performance by balancing the two. A higher F_1 score usually indicates that the system has struck a good balance in the matching process, ensuring both matching accuracy and covering more actual positive samples.

$$F_1 = 2 * (percision * recall)/(percision + recall)$$

Table 1: Accuracy, recall and F_1 for different data volumes.

Volume of data/ 10,000 entries	Precision	Recall	F_1
1	85.33%	96.37%	0.9051
5	88.04%	96.89%	0.9225
10	91.28%	97.42%	0.9425

4.3. Experimental validation

(1)Demonstration of Q&A system

In this paper, a Q&A system based on BERT model and knowledge graph is developed for the field of educational psychology. The system is developed using the Flask framework in Python, and the Neo4j graph database is chosen as the database based on its advantages in handling complex relationships and large-scale graph data. This system is based on the MVC (model-view-controller), a design idea commonly used in software design, which divides the system into a six-layer structure of data layer, knowledge graph construction, logic processing layer, user interaction layer, personalized learning support, and intelligent educational decision-making.

In order to ensure the compatibility of the system in various browsers and maintain a fast response, jsp + jquery front-end page development, to achieve the front-end and back-end data interaction. jsp provides dynamism and interactivity, embedded in the HTML to display dynamic content. The front-end interface is divided into the question area, answer area and feedback area. The question area allows users to input natural language questions, and the system presets recommended questioning methods to guide more accurate questioning. The answer area displays precise and concise search results, and provides answer-related named entities, similarity-matching question templates, and generated Cypher query statements, so that users can understand the system's query processing. The feedback area is used for users to evaluate the quality of answers and provide valuable references for further improvement of the system. The overall design makes it easy for users to input questions, obtain accurate answers and related information, and provide feedback, realizing an efficient, friendly and easy-to-use user experience.

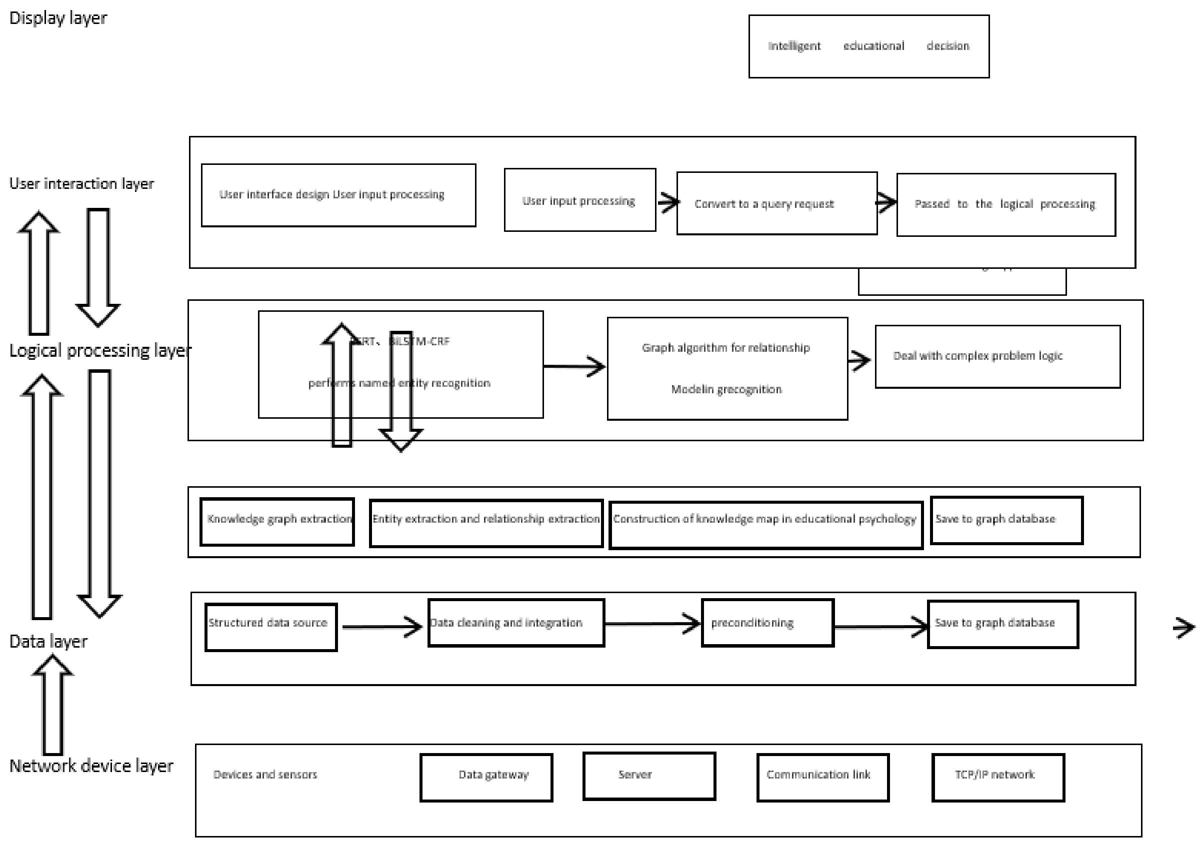

Figure 4: System Architecture Diagram.

(2) System performance analysis

In order to evaluate the accuracy of the quiz system, we designed 300 questions related to educational psychology under the guidance of experts as data for testing the quiz system. Then, the answers returned by the system were evaluated to test the performance of the Q&A system.

5. Conclusions

Knowledge graph-based Q&A system is one of the important means to apply knowledge graph to practical scenarios. This paper proposed a set of domain knowledge graph construction methods, including data preparation, text data labeling, knowledge extraction, knowledge storage, and other steps. For structured and unstructured data, this paper given the corresponding processing methods. In particular, the unstructured data exists a large number of redundancy data, which fails to be used to directly construct knowledge graph. Pre-processing of educational psychology knowledge point data before the experiment, and then combined with the professional knowledge in the field of educational psychology to extract the corresponding entity relationship characteristics of each knowledge point, and then the extracted entities are stored through the storage tool, and finally form the knowledge graph of educational psychology. The experiment results show that the effectiveness of the proposed knowledge graph-based Q&A system on educational psychology applications. The proposed knowledge graph-based Q&A system can achieve the average answer accuracy of 88.33% and provided an effective method for the application of knowledge graph in other subject areas. Our work not only provided an efficient and accurate Q&A system for the field of educational psychology, but also provided important support for the further development of knowledge graph in disciplinary research and application.

References

[1]Hartmann, W. K. (1975) Lunar "cataclysm": A misconception. Icarus, 24(2), 181-187.

[2]Green, B. F., Alice, K. W., Chomsky, N., et al. (1961) Baseball: an automatic question-answerer. Proceedings of the Western Joint Computer Conference. IEEE Computer Society.

[3]Weizenbaum, J. (1966) ELIZA-Acomputer Program For the Study of Natural Language Communication Between Manand Machine. Communications of the ACM, 9(1), 36-45.

[4]Divyatharshini, G., Devi, P. G., Srimathi, S., et al. (2018) An Efficient Community Question Answer (CQA) using the Metadata Rating. IJARIIT, 2.

[5]Uyar, A., Aliyu, F. M. (2015) Evaluatings earch features of Google Knowledge Graphand Bing Satori. Online In formation Review, 39(2), 197-213.

[6]Bollacker, K. (2008) Freebase: a collaboratively created graph database for structuring human knowledge. Proceedings of the 2008 ACMSIGMOD International Conference on Management of Data, 1247-1250.

[7]Vrandecic, D., Krtoetzsch, M. (2014) Wikidata: a free collaborative knowledgebase. Communications of the ACM, 57(10), 78-85.

[8]Suchanek, F. M., Kasneci, G., Weikum, G. (2007) YAGO: acoreofsemantic knowledge unifying Word Netand Wikipedia. Proceedings of the 16th International Conference on World Wide Web, 697.

[9]Auer, S., AI, E. (2007) DBpedia: anucleus for awebofopen data. Proceedings of the 6th Semantic Web, 722-735.

[10]Richens, R. H. (1958) Interlingual Machine Translation. The Computer Journal, 3, 144-147.

[11]Lenat, D. B. (1989) Building Large Knowledge-Based Systems–Representation and Inference in the CYC Project. Addison-Wesley Pub. Co.

[12]Marti, A. (1998) WordNet: An Electronic Lexical Database. Automated Discovery of Word Net Relations, 5.

[13]Kuck, G. (2004) TimBerners-Lee's Semantic Web. SA Journal of Information Management, 6(1).

[14]Vrandecic, D., Krtoetzsch, M. (2014) Wikidata: a free collaborative knowledgebase. Communications of the ACM, 57(10), 78-85.

[15]Hodges, A. (1984) Alan Turing: The Enigma. Physics Today, 37(11).