Journal of Cardiovascular Disease Research,
Volume 12 (4), ISSN: 0975-3583, 0976-2833

ORIGINAL ARTICLE

Google Bert- Multiple Choice Question Generation on Ontology Base

Gadi Himaja [#1], Sri Harshitha Gadu [#2], Kamisetty Venkata Harshith [#3], Muvvala Yamini [#4], Sundaraneedi Sai Sravya [#5], K.Vivek Murahari [#6], Sravanthi Mannava [#7]

[#1] Assistant Professor, Department of Computer Science and Engineering, Gitam Institute of Technology, Visakhapatnam-530045.

[#2, 3, 4, 5, 6, 7] Students, Department of Computer Science and Engineering, Gitam Institute of Technology, Visakhapatnam-530045.

Received: 15 March 2021; Accepted: 20 May 2021; Published: 20 July 2021

Abstract

Data mining is the ability to identify useful information from raw data and process the useful information in a separate list. The process of extracting useful information from big data source is known as ontology. In current days the semantic web is becoming a more trending topic for a lot of new research inventions. Multiple Choice Question Generation is one of the hot topics which are becoming a challenge for semantic web developers. By using this ontology model, a lot of researchers try to proposed MCQ item generation by giving sample phrases or sentences, or paragraphs for the system and try to generate the MCQ automatically based on those phrases. Hence in this proposed paper, we try to design an MCQ item generation system and we try to label that system as OntoQue. The proposed application is evaluated and tested using Deep Learning NLP pre-trained model like Google BERT, through which we can able to extract the MCQ generation for a summary. In the primitive methods, we try to use Long Short-Term Memory (**LSTM**) networks in order to learn order dependence in a sequence manner, but this is failed to achieve 100 % accuracy in training bulk amount of data and there is also a problem in LSTM like vanishing gradient problem through which some quantity of data vanishes while training the machine. For testing the proposed application we try to use the Stanford question answering dataset, in which the data set combines more than 1one lakh questions in SQuAD1.1 with over fifty five thousand unanswerable questions written adversarial by crowd workers to look similar to answerable ones. The proposed model can accurately train with this dataset and this can try to generate the desired result for any type of text summary which we try to enter for the system.

Keywords:

Long Short-Term Memory, Deep Learning, Google BERT, Ontology Model

---

Journal of Cardiovascular Disease Research,
Volume 12 (4), ISSN: 0975-3583, 0976-2833

1. Introduction

Ontologies are knowledge representation models that provide a very good platform for developing interactive applications. Nowadays a lot of advancements came in the web and one among them is the semantic web. This semantic web has constructed and gained more and more interest among researchers for designing ontology-based applications in numerous research areas. One of the researches in the field of the semantic web is question generation (QG), a subfield of artificial intelligence [3]-[5]. In current days for all the researchers, ontology-based multiple choice question (MCQ) generation is becoming a challenging task. This MCQ generation is proved to be an efficient tool for measuring the achievement of learners. Instructors could benefit from such automatic MCQ generation systems [1] for conducting tests automatically, since the task of manually constructing MCQ items for tests is very complex and time-consuming. Also, the manual construction of MCQ items doesn't give high-quality result for that corresponding phrase. Even though ontology-based MCQ generation systems successfully generate MCQ for the given phrase, a lot of pending work or research is left in order to evaluate how well these MCQ items need to be deployed in an educational system or online examinations [2]. A keen evaluation and observation is required for evaluating the guidelines and adjust requirements for the design and development of ontology-based MCQ generation systems.

From the below figure 1, we can clearly represent an Ontology model for classifying the MCQ for the students based on three levels:

1. Expert
2. Intermediate
3. Beginner

From the given phrase, the MCQs need to be classified into three different categories based on the type of category which is opted by the concerned person. The educator will give options for each and every student to choose the level while they start the exam. Once the type is chosen now the MCQ generation system will automatically generate several MCQs and their corresponding answers based on type. If any student is an expert in that topic, he can choose the expert type, and immediately the MCQ generation system will generate all the questions and answers related to the expert level. If the same student who is having very little knowledge about that concept, he can choose intermediate level type so that the MCQ generation system will generate MCQs based on the intermediate level. If the same student chooses the ability type as a beginner then the MCQ generation system will generate the MCQs at a very beginner level[6].

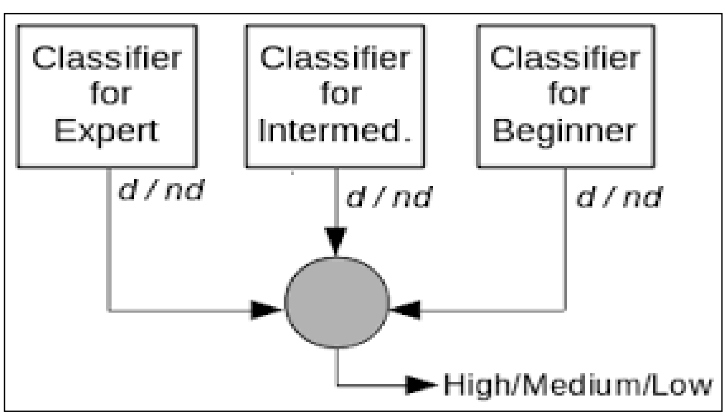

Figure. 1. Represents the Sample MCQ generation Scenario on Online Examination

This proposed paper mainly aims to address this issue by assessing the performance of these systems in terms of the efficiency of the generated MCQs and their pedagogical value. The primitive models are not completely successful in designing an MCQ generation based on text summary in a dynamic manner without vanishing the data which is given for the system. In general the existing system we try to use Long Short-Term Memory (LSTM) networks in order to learn order dependence in a sequence manner, but this is failed to achieve 100 % accuracy in training huge amount of data and there is also a problem in LSTM like vanishing gradient problem through which some quantity of data has vanished while training the machine [7]. Hence the proposed Google BERT is having no problem and this can hold millions of words in order to check the summary and generate the Question and Answers along with MCQs dynamically. We present an experimental evaluation of an ontology-based tool for generating MCQ items, the system known as OntoQue[8]-[10]. OntoQue is a question generation system that assists the instructor by automatically generating assessment items using domain ontology.

2. LITERATURE SURVEY

Literature survey is that the most vital step in the software development process. Before developing the new application or model, it's necessary to work out the time factor, economy, and company strength. Once all these factors are confirmed and got approval then we can start building the application. The literature survey is one that mainly deals with all the previous work which is done by several users and what are the advantages and limitations of those previous models. This literature survey is mainly used for identifying the list of resources to construct this proposed application.

MOTIVATION

1) Automatic Generation of Multi-Modal Dialogue from Text Based on Discourse Structure Analysis. Published In Proceedings 1st IEEE International Conference on Semantic Computing (ICSC- 07), 2007, pp. 27–36.

AUTHORS: H. Prendinger, P. Piwek, and M. Ishizuka

In this paper the authors mainly classify the automatic question generation from text into two main categories

a) Dialogue and interactive Q/A system

b) Education assessment.

The proposed research work is very interesting and for the current application, the researchers try to generate the questions and answers for the expository texts. But the results didn't derive the accurate reports and left some incompleteness in generating the accurate MCQ and its answers.

2) Experiments with Interactive Question-Answering", Published In Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics, 2005, pp. 205–214.

AUTHORS: S. Harabagiu and D. Moldovan

In this paper the authors mainly concentrated on the MCQ generation-related domain, the authors mainly concentrated on predicting MCQ generation based on user wishes. Here the authors try to generate the MCQ based on the user wish and the user wishes to ask. In this, they compared with some other paper which tries to identify the domain knowledge first and then find out the context related to that domain. By comparing that paper the authors want to generate the MCQs based on user topic, the comparative results clearly state that the proposed MCQ generation system is very accurate in deriving those MCQs but it is not efficient if an unknown domain is trained by the system. If the user enters a new domain phrase and asks for the MCQ generation the system may not generate accurate results because the system doesn't have any information related to that new domain.

3) Automatic Question Generation from Text – an Aid to Independent Study SIGCUE Outlook", Published In proceedings of the SIGCSE-SIGCUE technical symposium on Computer science education, 1976, pp.104–112.

AUTHORS: J. H. Wolfe

In this paper, the authors mainly discussed the context of generating questions for education assessment, which is under investigation for many years. In fact, create an exam paper is a very complex task and the generation of MCQs for the educational system is very time-consuming and hence the instructors require a lot of workload in order to generate MCQs for the educational systems. In this paper the authors gone with deep research about the educational assessment system and then created a model for Multichoice questions for text comprehension. The main aim for the current authors is to design an application that can automatically generate the MCQs from the text, independent of domain knowledge.

3. EXISTING SYSTEM AND ITS LIMITATIONS

In the existing system, there was no proper method to identify the words from the summary and find out the question, answers and generate the best MCQs for that input string. All the existing approaches are best in training text data which is smaller in size and accurate in generating the MCQ from that pre-defined trained data. But no method is having the ability to test the text summary dynamically based on any topic despite taking pre-defined summary. The following are the main limitations of the existing system.

LIMITATIONS OF THE EXISTING SYSTEM

1. Inaccurate Results
2. There is a vanishing gradient problem in the existing methods
3. The existing method LSTM is not efficient to generate MCQ for large data.
4. This is not accurate to test on dynamic inputs that are train data and test strings with different data are not supported in the LSTM.
5. There is a lot of time complexity to generate the MCQ for the text summary.

4. PROPOSED SYSTEM AND ITS ADVANTAGES

The proposed Google BERT is having no problem and this can hold millions of words in order to check the summary and generate the Question and Answers along with MCQs dynamically. For testing the proposed application we try to use the Stanford question answering dataset, in which the data set combines more than 100,000 questions in SQuAD1.1 with over 50,000 unanswerable questions are written adversarial by crowd workers to look similar to answerable ones. The proposed model can accurately train with this dataset and this can try to generate the desired result for any type of text summary which we try to enter for the system. The following are the advantages of the proposed model:

ADVANTAGES OF THE PROPOSED SYSTEM

1) By using the proposed deep learning model Google BERT we can able to generate MCQs accurately for large data.
2) It generates a very accurate result
3) It is less time complexity
4) There is no vanishing gradient problem in the proposed Google Bert Model.
5) This is very efficient in searching dynamic summaries I.e Test and Train data may not be the same.

5. PROPOSED MCQ GENERATION USING GOOGLE BERT MODEL

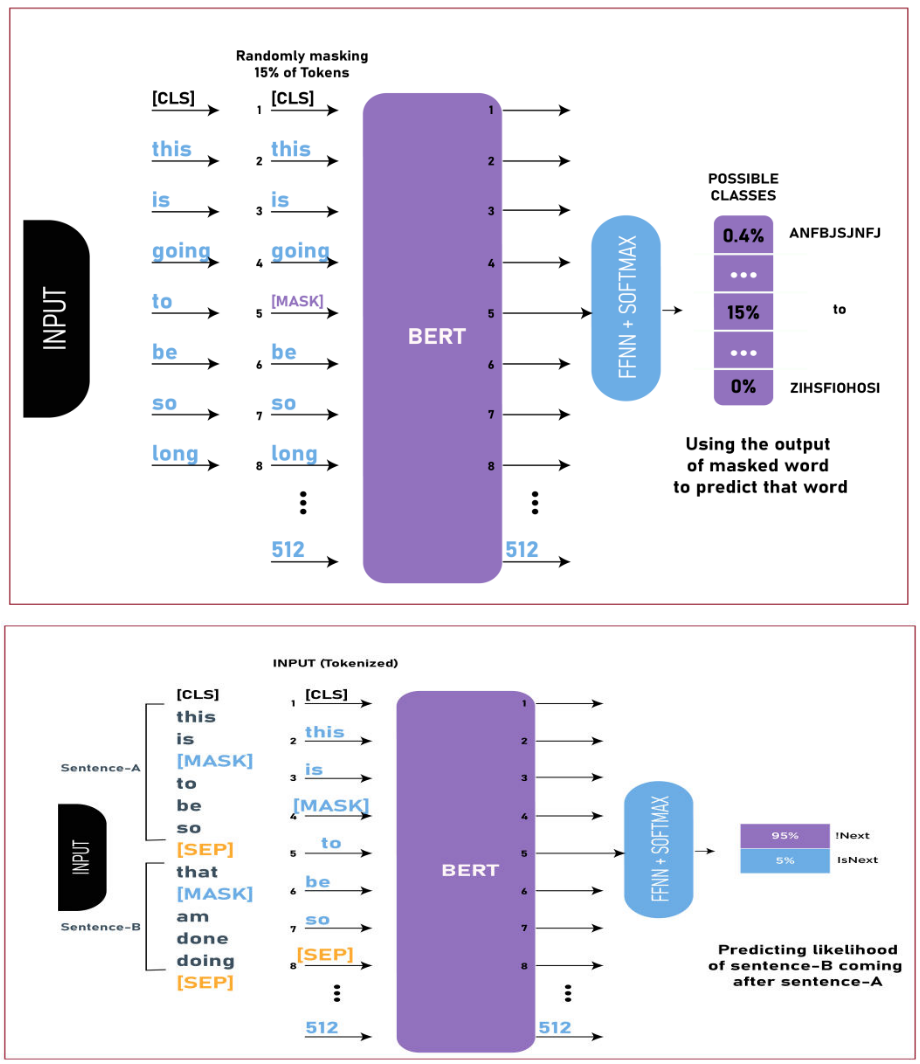

Figure. 2. Represents the Proposed MCQ Generation System Using Google BERT

BERT stands for Bidirectional Representation for Transformers. This was developed by researchers at Google Research in 2018. This is mainly designed in order to increase the understanding of the Google Search Queries. In general, google try to encounter fifteen percent of new queries and try to find out answers for those new queries which are generated on a day. Hence there is a great need of designing a new system that can encounter all-new queries and try to understand the queries easily and generate meanings for them. From the above figure 2, we can clearly identify the MCQ generation system on a text phrase or sentence [11]. Initially, we try to choose to input a text or sentence with a set of meaningful conversations. The google BERT follows the natural language processing (NLP) technique in order to divide the message into parts and identify the meaning of each and every word which is spat. In general, all the split words are assigned with Tokens (T) and for each and every token google BERT conducts research and identifies the meaning of each and every word. During the research, some words are treated as not important and hence such words need to be masked and they are labeled as MASK Token. By using this BERT model, we try to identify all such MASK tokens and try to improve the model by applying the fine-tuning technique on that MASK tokens. Here we try to calculate the probability of output by using a fully connected layer also known as the Softmax layer. From figure 2(a) we can see the message is divided into parts and we can see 85 % of data is tokens generated and they are Non-Mask tokens and 15 % of data is MASKED tokens. Hence we primarily identify the loss on those 15 % masked words from the given input text [12].

In the next phase, we try to apply the BERT technique to those two separate sentences which are divided in the previous stage and our main goal is to predict whether the second sentence is next subsequent sent to the primary one or whether this should be removed. Here we take fifty percent of data in a single attempt and then train with the BERT model and then calculated next sentences randomly using a label is next and if it is not coming as next sentence we label it as NotNext. Since this model is a classification model, we first try to take CLS token which is a class token for comparing the two sentences whether they belong to a single class or not. If the CLS token gets the same value then this will use the help of the SEP token to separate the two sentences and then join those sentences in a meaningful manner to form a single sentence. From the above figure 2(b) we can clearly see that the BERT model clearly predicted the sentence B is likelihood coming after sentence A.

6. BERT ARCHITECTURE

BERT is a multi-layered encoder is a two stage method which uses two models like BERT base and BERT large. The BERT base model is mainly consisting 12 layers with 110 million parameters, when compared to BERT large it is having 24 layers with 340 million parameters.BERT base – 12 layers, 12 attention heads, and 110 million parameters.BERT Large – 24 layers, 16 attention heads and, 340 million parameters.

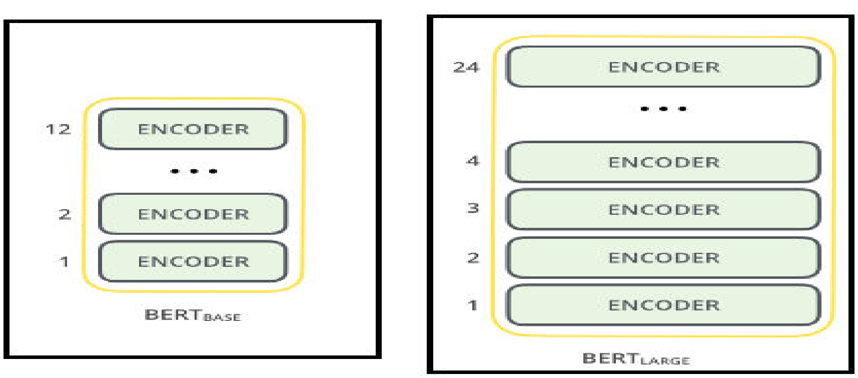

Figure. 3. Represents the Proposed BERT Architectures

Journal of Cardiovascular Disease Research
Journal of Cardiovascular Disease Research,
Volume 12 (4), ISSN: 0975-3583, 0976-2833

COMPARISON OF BERT WITH SOME PRE-TRAINED MODELS IN THIS GIVEN CONTEXT

	BERT	RoBERTa	DistilBERT	XLNet
Size (millions)	**Base**: 110 **Large**: 340	**Base**: 110 **Large**: 340	**Base**: 66	**Base**: ~110 **Large**: ~340
Training Time	**Base**: 8 x V100 x 12 days* **Large**: 64 TPU Chips x 4 days (or 280 x V100 x 1 days*)	**Large**: 1024 x V100 x 1 day; 4-5 times more than BERT.	**Base**: 8 x V100 x 3.5 days; 4 times less than BERT.	**Large**: 512 TPU Chips x 2.5 days; 5 times more than BERT.
Performance	Outperforms state-of-the-art in Oct 2018	2-20% improvement over BERT	3% degradation from BERT	2-15% improvement over BERT
Data	16 GB BERT data (Books Corpus + Wikipedia). 3.3 Billion words.	160 GB (16 GB BERT data + 144 GB additional)	16 GB BERT data. 3.3 Billion words.	**Base**: 16 GB BERT data **Large**: 113 GB (16 GB BERT data + 97 GB additional). 33 Billion words.
Method	BERT (Bidirectional Transformer with MLM and NSP)	BERT without NSP**	BERT Distillation	Bidirectional Transformer with Permutation based modeling

From the above table, we can clearly identify BERT is having more performance compared with several other models when we compare multiple models with same performance inputs.

7. EXPERIMENTAL RESULTS

Implementation is a stage where the theoretical design is converted into a programmatic manner. In this proposed application we try to use PYTHON as a programming language in which Google Collaboratory is used as a working platform to text the current application.

1) IMPORTING ALL NECESSARY LIBRARIES

```
!pip install gensim
!pip install git+https://github.com/boudinfl/pke.git
!python -m spacy download en
!pip install bert-extractive-summarizer --upgrade --force-reinstall
# !pip install spacy==2.1.3 --upgrade --force-reinstall
!pip install -U spacy
!pip install -U nltk
!pip install -U pywsd
#Python Implementations of Word Sense Disambiguation
#FlashText is a Python library created specifically for the purpose of searching
!pip install flashtext
import nltk
nltk.download('stopwords')
nltk.download('popular')
```

From the above pseudo code we can see all the libraries which are required for MCQ generation is imported and we can see BERT and natural Language Tool Kit (NLTK) are also imported. Here from the NLTK library we are going to import two main methods like Stop words and popular, which are two efficient methods to remove the stop words and find out the most useful and popular words which are present in the given phrase. Here BERT library is used for summarizing the text from the given phrase or sentence.

2) **USER TRY TO IMPORT THE INPUT DATA**

Importing text data

```
from google.colab import files
files.upload()
```

```
Choose Files No file chosen        Upload widget is only available when the cell has been executed in the current browser session.
Please rerun this cell to enable.
Saving Greeks.txt to Greeks (1).txt
{'Greeks.txt': b'The Greek historian knew what he was talking about. The Nile River fed Egyptian civilization f
```

Figure . Represents the User try to import the text data

From the above figure we can clearly see that user try to browse a text data with name Greeks.txt from the Desktop. Now the file is uploaded into the drive and now we can able to read the content present in that input file.

```
!ls
```

```
'Greeks (1).txt'    sample_data
```

Now we can see the file is loaded successful

3) **BERT SUMMARIZER IS APPLIED**

```
from summarizer import Summarizer
f = open("Greeks.txt","r") # enter the name of file to read
full_text = f.read()
model = Summarizer()
result = model(full_text, min_length=60, max_length = 500 , ratio = 0.4)
summarized_text = ''.join(result)
summarized_text
```

Here we try to apply the BERT summarizer method and find out the MCQs for that given phrase

4) TEXT IS DIVIDED INTO PARTS

```
['The Nile River fed Egyptian civilization for hundreds of years.',
 'It begins near the equator in Africa and flows north to the Mediterranean Sea.',
 'A delta is an area near a river’s mouth where the water deposits fine soil called silt.',
 'This soil was fertile, which means it was good for growing crops.',
 'The red land was the barren desert beyond the fertile region.',
 'When the birds arrived, the annual flood waters would soon follow.',
 'Then they used a tool called a shaduf to spread the water across the fields.',
 'These innovative, or new, techniques gave them more farmland.',
 'They were the first to grind wheat into flour and to mix the flour with yeast and water to make dough rise in
 'Egyptians often painted walls white to reflect the blazing heat.',
 'Poorer Egyptians simply went to the roof to cool off after sunset.',
 'Even during the cool season, chipping minerals out of the rock was miserable work.',
 'One ancient painting even shows a man ready to hit a catfish with a wooden hammer.',
 'A boomerang is a curved stick that returns to the person who threw it.)',
 'The river’s current was slow, so boaters used paddles to go faster when they traveled north with the current.
```

From the above figure we can clearly see that input data is divided into meaningful sentences and now we can generate the MCQs and their corresponding options from this input phrase.

5) MCQ GENERATION FROM TEXT

```
###############################################################################
                    MCQ GENERATION FROM TEXT
###############################################################################


1) The Nile provided so well for _______ that sometimes they had surpluses, or more goods than they needed.
         a )   Egyptians
         b )   Algerian
         c )   Angolan
         d )   Bantu

More options:  ['Basotho', 'Beninese', 'Berber', 'Black African', 'Burundian', 'Cameroonian', 'Carthaginian', 'Chadian', 'Chewa', 'Congolese

correct answers is : egyptians
```

```
2) As in many ancient societies, much of the knowledge of _______ came about as priests studied the world to find ways to please the gods.
         a )   Kuwait
         b )   Iraq
         c )   Egypt
         d )   Saudi Arabia

More options:  ['Jordan', 'Israel', 'Fertile Crescent', 'Turkey', 'Iran', 'Lebanon', 'Shari', 'Mauritania', 'Nigeria', 'Somali peninsula', '

correct answers is : egypt

3) The _______ provided so well for Egyptians that sometimes they had surpluses, or more goods than they needed.
         a )   Entebbe
         b )   Gulu
         c )   Nile
         d )   Buganda

More options:  ['Jinja', 'Lake Edward', 'kayunga', 'gulu', 'entebbe', 'Port Sudan', 'Omdurman', 'Darfur', 'Libyan Desert', 'Kordofan', 'Khar
```

From the above figure we can see MCQs are generated from a given input. Here we can see questions along with four options and more options are suggestions. Here at the end, we can see a clear answer for that generated MCQ.

8. CONCLUSION

In this current work we for the first time designed and implemented an MCQ (Multiple Choice Question Generation) model which is one of the hot topics for semantic web developers. In this proposed paper we try to design an ontology-based MCQ item generation system using google BERT known as OntoQue. The proposed application is evaluated and tested using Deep Learning NLP pre-trained model like Google BERT, through which we can be used to extract the MCQ generation for a summary. In the primitive methods, we try to use Long Short-Term Memory (**LSTM**) networks in order to learn order dependence in a sequence manner, but this is failed to achieve 100 % accuracy in training bulk amount of data and there is also a problem in LSTM like vanishing gradient problem through which some quantity of data has vanished while training the machine. By conducting various experiments on our proposed model, by taking some phrases from Stanford University, we try to apply the google BERT model and generate the MCQ for that input data.

9. REFERENCE

[1] A well-known author Manish Agarwal and Rakshit shah et.al, has written a paper on "Automatic Cloze-Questions Generation by Annamaneni Narendra", published in LTRC, 2016.

[2] A well-known author Manish Agarwal and Prashanth Mannem et.al, has written a paper on "Automatic Question Generation using Discourse Cues", published in Proceedings of the Sixth Workshop on Innovative Use of NLP,2011.

[3] A well-known author C. Rajeswari et.al, has written a paper on "Automatic Question Generation Using Software Agents for Technical Institutions", published in International Journal of Advanced Computer Research, December-2013.

[4] A well-known author Ming Liu and Rafael A. Calvo et.al, has written a paper on" G-Asks: An Intelligent Automatic Question Generation System for Academic Writing Support", Published with DOI: 10.5087/dad.2012.205,2012.

[5] A well-known author I.E. Fattah et.al, has written a paper on "Automatic Multiple Choice Question Generation System for Semantic Attributes Using String Similarity Measures", Published in www.iiste.org,2014.

[6] A well-known author Fattoh et.al, has written a paper on "Semantic-Based Automatic Question Generation using Artificial Immune System", Published in Computer Engineering and Intelligent Systems,2014.

[7] A well-known author Y.SKALBAN et.al, has written a paper on "Automatic question generation in multimedia-based learning", published in https://www.aclweb.org/anthology/C12-2112.pdf.

[8] A well-known author A.E.Awad et.al, has written a paper on "Automatic Generation of Question Bank Based on Pre-defined Templates", Published in IJIACS, April 2014.

[9] A well-known author Andreas et.al, has written a paper on "Automatic Generation Of Multiple Choice Questions From Domain Ontologies", Published in IADIS,2008.

[10] A well-known author Payal Garg and Er. Charandeep Singh et.al, has written a paper on " A Review on Question Generation System From Punjabi Text", Published in IJETTCS 2014.

[11] A well-known author R.C.Samant et.al, has written a paper on "An Overview of Automatic Question Generation Systems", Published in IJSETR, 2014.

[12] A well-known author Lee Becker et.al, has written a paper on "Mind the Gap: Learning to Choose Gaps for Question Generation", published at https://dl.acm.org/doi/10.5555/2382029.2382150.