OntoQuest: An Ontological Strategy for Automatic Question Generation for e-assessment using Static and Dynamic Knowledge

[1]Gerard Deepak, [2]Naresh Kumar, [2]G VSN Sai Yashaswea Bharadwaj, [1]A Santhanavijayan
[1]Department of Computer Science and Engineering
[2]Department of Mechanical Engineering
[1,2]National Institute of Technology, Tiruchirappalli.
111116017@nitt.edu

***Abstract* – In modern times, e-learning has become the most convenient and significant environment and is being used to attain knowledge about and across various domains. While there are plenty of e-learning resources available across the ever-expanding web, imbibing knowledge is not the only factor that contributes to the edification of an individual. The assessment also must be given a quintessential role because both learning and assessment are two sides of a coin. In this paper, OntoQuest, which is a scheme for generation of multiple-choice questions is done based on the domain or subject of choice from the user. An equivalence set based summarization scheme has been proposed. Synonymization and Anonymization are used for the generation of best key and distractors respectively. Domain and Granular Ontologies are crawled for relevant sub-topics and auxiliary topics. WordNet integrates dynamic knowledge for improving the overall accuracy. Jaccard similarity is encompassed for the generation of the correct key. OntoQuest is a reliable framework that yields an overall accuracy of 95.46 % and 95.05% for key and distractor generation respectively, which is higher and consistent compared to the existing models available.**

Keywords— Domain Ontology, e-assessment, e-learning, Granular Ontology, Question generation

I. INTRODUCTION

Education without assessment is merely an incomplete process of acquiring knowledge. The process of learning must always be accompanied by the simultaneous and consistent process of assessment in order to keep track of learner's progress and thereby abetting improvement at various stages of this learning process. A series of reliable assessments are the indicators of profundity attained by learners and provides an opportunity for a healthy competition among peers, which ultimately facilitates a crescendo and makes the entire process comprehensive. A reliable assessment strategy must be able to test the learners' ability to grasp various attributes of the entire given domain. The questions that make up the complete assessment must enhance logic and contextual reasoning. The respective answers must efficiently analyze and produce accurate explanations. When these questions are reasoned with cognizance, and the corresponding answers are correctly marked, the process of learning can be considered to nearly culminating.

Multiple Choice Questions are the most common among the various types of questions that prevail. These questions require vigilance, comprehension about the given topic and also require scrutiny along with reasoning, which often comes into play during option elimination. These questions prevail in various Standard and International examinations, mainly because distractors always accompany these questions. There is only a 20% probability in guessing the right answer among all the available five options. Only when the knowledge is thoroughly attained, the correct option can be confidently marked. Thus, these Multiple-Choice Questions are very reliable for an assessment. The questions must be able to indicate the context clearly and not merely repeat the sentences present in the corpus. The process of modeling a proper methodology to generate Multiple Choice Questions is not complete without generating a good set of distractors. These distractors play a pivotal role because they test the reasoning capacity of the learner by providing, nearly close but not correct answer choices. These nuances between the options, when clearly scrutinized by the student will pave the way to a fruitful learning experience.

Albeit the advances in the fields of NLP, generation of distractors is quite a tedious task. This paper proposes a novel strategy based on ontologies, designed to give proper weighing to the synonyms and antonyms of the answer key and use these are potential distractors based on the semantic score. This model provides a dynamic approach to the existing problem, since, the formulated ontology can always be updated in accordance with the changes happening across the globe. The distractors generated should have some salient features so that it should exactly mean the same as the answer key, which invigorates a sense of ambiguity to the test taker. A perfect distractor must be very close to the answer key while having a considerable difference between the same. This process needs to be meticulously taken care of as it is the essence of the entire question generation process.

Motivation: With a pedagogical and paradigm shift in the conduct of several online examinations using Multiple Choice Questions, it is highly cumbersome to manually frame such questions. It is essential for an accurate and automated system for subject adherent e-assessment. the world of ever-growing knowledge and data, it becomes quite mandatory for any Teaching institution to adapt to the fast pacing online world and be equipped with the sophisticated models of generating proper questions to any given domain so that the true ingenuity of a student is thoroughly monitored. This provides motivation for this paper and kindles the need for a relevant and consistent approach to arrive at the generation of Multiple-Choice Questions to facilitate the edification of a student.

Contribution: An innovative framework for the automatic generation of Multiple-Choice Questions for e-assessment has been proposed. A novel Strategic summarization algorithm based on POS-tagged Equivalence set partition and term frequency has been proposed. The other novel contributions include the use of Domain Ontologies, Ontology Granularization and a Semantic Similarity based Key and distractor generation has been proposed. The supplication of auxiliary dynamic knowledge in the premise of static ontological knowledge is an added novelty in the framework. The precision, recall, accuracy, and f-measure for summarization, key and distractor generation have been increased.

Organization: This remaining paper is organized as follows. Section II depicts the related work. Proposed Architecture is described in Section III. Implementation is described in Section IV. Performance analysis and results are depicted in Section V. The paper is concluded in Section VI.

II. RELATED WORK

Boyuan Pan et al., [1] has proposed a model task named Conversational Question Generation (CQG) for conversation and passage-based question generation. The proposed model was able to drive a question-answering style conversation. An approach named ReDR is proposed and CoQA dataset has been used for the implementation. Beason et al., [2] put forward a computer-implemented model for the creation of question-answer pairs. The proposed model leverages domain-specific resources including lexicons, glossaries. This model also eliminates erroneous pairs generated. Santhanavijayan et al., [3] have proposed a model for multi swarm optimization based automatic ontology for e-assessment. The model encompasses the extraction of text from the web utilizing the UQR algorithm. Summarization is then performed on the text and the MSO method is incorporated into the methodology.

Diatta et al., [4] proposed a model for a bilingual ontology helping learners various and enable them to self- evaluate various laboratory material and safety notes. Leo et al., [5] proposed a model focussed on concentrating on the limitations present in various approaches for question generation. An ontology-based is incorporated for the implementation. Specific case-based questions have been formulated which include assembling complex stems, suitable option selection, provision of explanations of option correctness and those of incorrect. The implementation is done on a medical ontology. Medical education is considered for the application of the model.

Chris et al., [6] have proposed an innovative strategy for the generation of question answering corpus. The approach strategically fuses both question generation as well as extracting answers such that roundtrip consistency is preserved. Santhanavijayan et al., [7] have proposed a Fuzzy MCS strategy for automatic question generation which is hybridized with a modified cuckoo search algorithm. The strategy uses ontologies but however, there is a scope to ontologically improve the overall performance by also the elimination of optimization algorithms for making it computationally less complex. In [8-17], there is an effective use of ontologies and semantic strategies for various scenarios which can definitely be employed for question generation schemes.

III. PROPOSED ARCHITECTURE

The architecture of the proposed OntoQuest framework for automatic question generation is depicted in Figure 1. OntoQuest employs a novel strategy for e-assessment by generating multiple-choice questions from various crawled web corpus, which takes the domain as prospective input from the user. The domain of choice and it’s associated sub-sub domains are taken as input, which is processed as topics and auxiliary topics from the existing Domain Ontology. These topics are further looked up in the web corpus and entire Domain Ontology is parsed to load the sub-concepts. The documents containing all the relevant Sub-Topics and Auxiliary topics that are generated are then loaded using the TF-IDF measure. Furthermore, once the documents comprising of relevant content are crawled, pre-processing and summarization are realized.

The pre-processing is achieved using the subsequent processes. Each of the text document from the document corpus is tokenized based on a customized Blank Space and Special Character tokenizer. Subsequently, Lemmatization is then performed where the base form of the word is derived using a WordNet Lemmatization strategy. Once the document is subjected to Tokenization and Lemmatization, Parts of Speech Tagging is achieved using a standard POS Tagger. These three steps of pre-processing enable further implementation of the devised model. The strategy requires the Summarization of each document, which is not merely a process of diminishing the number of sentences available. The strategy of summarization must be in such a way that the essence of any given text corpus must not be lost and its preservation is a mandate. The summarized sentences are the ones, the questions will most likely be based on. Henceforth, after summarization, important sentences serve as markers for generating questions.

To achieve summarization, an equivalent set of Proper Nouns (P_N), Adjectives (A_j), and Adverbs (A_v) is made use of. For all these words, based on the Term Frequency, a separate equivalent set of frequent words is generated as F_W. All the sentences containing these Frequent words are concatenated in the set F_S. For all the sentences in this set, Proper Nouns appearing at the successive sentences are removed and the document is concatenated into a single Summarized file. For, Generation of the questions, each sentence with an identified keyword is parsed and is used for Question formulation, by appending keywords like why. what, define, append, list, appraise, etc, using standard templates obtained from a trained Bloom’s Taxonomy corpus. Further, on arriving at a standard question, key generation by either using the same keyword or further synonymizing of the same is followed. If synonymizing is the strategy, then the synonyms are generated from the domain ontology of a specific subject which is further granularized by using existing ontologies and WordNet. The synonyms are used as keys.

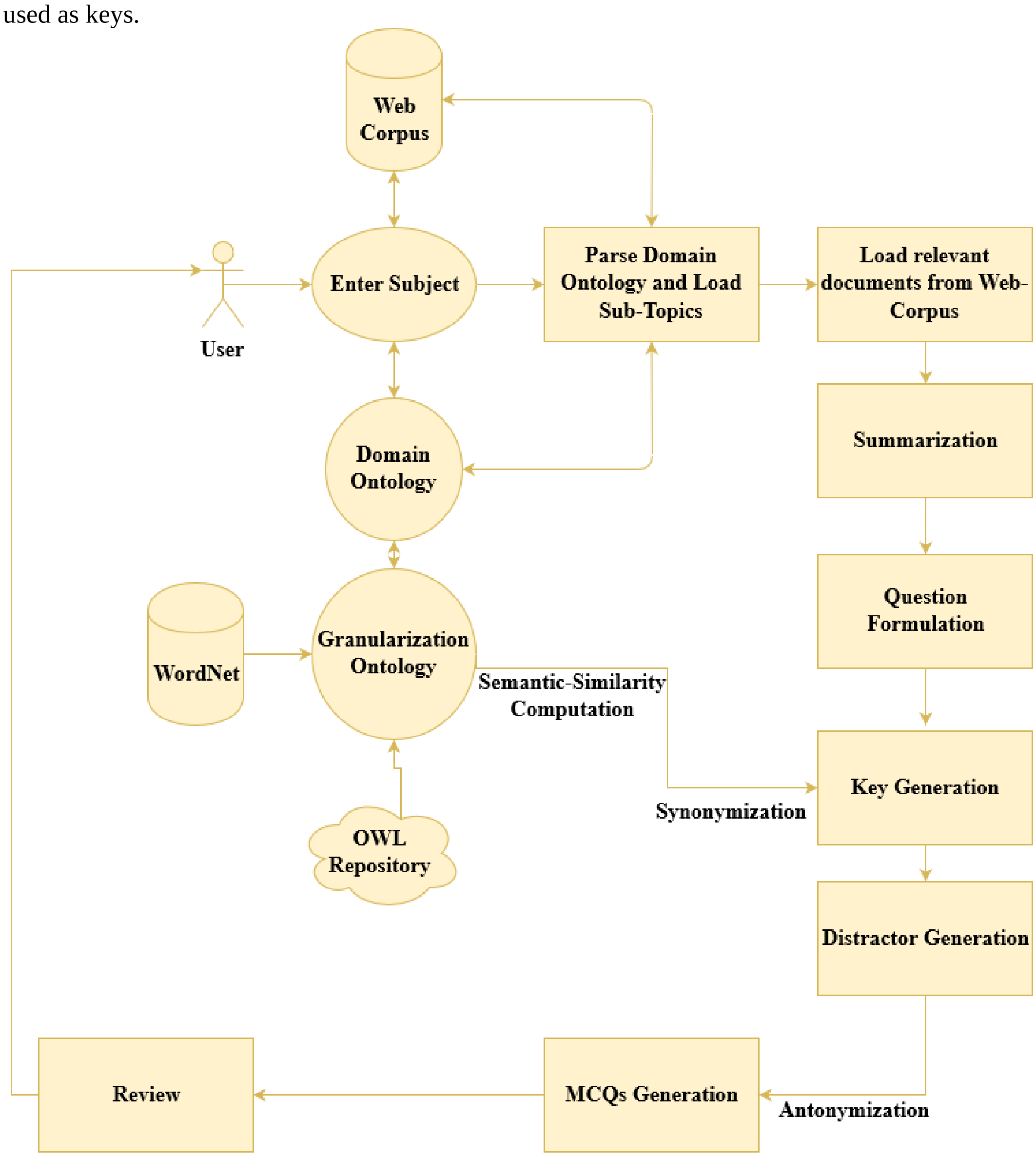

Figure 1: System Architecture of OntoQuest

Once key generation has been successfully completed, it is further succeeded by distractor generation, which is achieved through anonymization, which is also achieved by using a combination of granular ontologies and WordNet. However, the granular ontologies are collated from the existing ontologies and verified domain ontologies. For the generation of these distractors, a dynamic ontology modeling is devised and is frequently updated with synonyms and antonyms. This enables a cohesive approach and the synonyms and antonyms together generate distractors. This approach makes the process of distractor generation very simple and efficient. Finally, once the questions are generated along with the key terms and MCQs, a random function is appended for option shuffling and is presented to the user for review before the questions are finalized and arranged into a PDF.

IV. IMPLEMENTATION

The implementation of the methodology is done in accordance with the proposed architecture. The sentences are initially tokenized and then lemmatized. The sentences that form the core of the given text are filtered out. Then proper questioned are formulated with the corresponding set of distractors. Topics are the terms that are in direct coherence with the domain. Sub-topics have associations with the topics. However, those sub-topic terms do not directly correlate with the domain. Auxiliary topics are those terms that can be identified relevant to the domain but has no direct correlation. The proposed Ontology-driven Framework question generation is implemented in JAVA on intel i7 processor with 16 GB ram. The Domain Ontology has been modeled using protégé. Ontology granularization has been achieved by linking common-sense ontology, sumo, Linguistic English language ontology along with linked open data. The web corpus has been crawled using a customized JAVA crawler. DEIXTO has been employed for aligning all the constituent ontologies.

Table 1: Proposed Algorithm for Summarization

Input: Text Corpus of the given Domain
Output: Summarized text document.
begin
Step 1. Obtain the subject from the user.
Step 2. Look up onto the Domain Ontology (D_o) to obtain sub and auxiliary topics of the given subject.
Step 3. Using TF-IDF measure load relevant documents of the given subject using Topic (t), Sub-Topic (st) and Auxiliary Topic ($a_x t$) measure.
Step 4. ***Summarization:***
(i)Pre-process the text by removing all the stop words in the sentences.
(ii)Tokenize the entire text into sentences and obtain the respective sentences.
(iii)Word Tokenize all the words in these sentences and POS tag the obtained words.
(iv)Formulate sets of Proper Nouns (P_N), Adverbs (A_v), and Adjectives (A_j) from the POS tagged sentences.
(v)for words (W) in {PN, A_v, A_j }
if TF(W)>(POS tagged words/Total words in the current document-Stop words).
(vi)Generate a set of frequent words by appending the obtained words in (v) into $F_{W.}$
(vii)Formulate the sentences containing these frequent words into F_S.
(viii)for each sentence (S) in F_S
if ((S does not contain P_N and A_V) &(Score>Length (S/Total words between periods containing the Frequent word (W).
(ix)Append the sentences obtained in (viii) into S_D.
(x)for each sentence (S) in S_D
Concatenate to the summary (S_u)
end

Table 1 and Table 2. depict the algorithms for summarization and Question, distractor generation respectively. The entire proposed architecture is modeled with these algorithms at the crux. Summarization is carried out effectively after obtaining relevant topics, Sub-topics and auxiliary topics using TF-IDF measure. As depicted in the above algorithm, the relevant text is pre-processed and then various words with their respective parts of speeches are obtained. Once the equivalent sets of Proper Nouns, Adjectives, Adverbs are obtained, using Term Frequency measure, frequent words are filtered out. Subsequently, the sentences containing these frequent words are segregated. These sentences are then worked on so that the gist of the entire text is preserved. Proper Nouns appearing at the later parts of the sentence are then removed and the summary is concatenated into a single text file.

Once the summarization process is finished, the successive process is the generation of appropriate questions and distractors. For this, all the sentences containing pivotal adjectives, Adverbs, and Proper nouns are considered blanks are generated at the appropriate position and Multiple-Choice Questions are generated. Key generation and distractor generation plays the most crucial part of the whole algorithm. To achieve this process, a Granular Ontology (GO) is used. Based on the Jaccard Similarity measure, Synonyms and Antonyms are generated. Jaccard Similarity is calculated between Text and the Granular Ontology. The threshold is set to 0.75 and 0.25. If the similarity is greater than 0.75, the words in the appropriate text are appended into Synonyms set. The words with similarity less than 0.25 and considered to be Antonyms and are merged into a different Antonyms set.

Table 2: Proposed Algorithm for Questions and Distractor Generation

Input: Summarized document (S_u)
Output: Generation of relevant questions and their respective distractors.
begin
Step 1. Generate the Equivalent sets of Adjectives (A_j), Proper Nouns (P_N), and Adverbs (A_v)
Step 2. for each Adjective in A_j and A_j!=NULL
for each sentence S_i in S_u (i=1,2,3,4…)
Generate blank at appropriate position.
Step 3. for each Adverb in A_v and A_v!=NULL
for each sentence S_i in S_u (i=1,2,3,4…)
Generate blank at appropriate position.
Step 4. for each Proper Noun in P_N and P_N!=NULL
for each sentence S_i in S_u (i=1,2,3,4…)
Generate blank at appropriate position.
Step 5. Key generation, using Jaccard Similarity (Jsim) between the text T_i and Granular ontology (GO)
Step 6. Initialize the equivalent sets Synonyms (S_y[]) and Antonyms (A_y[])
Step 7. if Jsim(T_i, GO) > 0.75
S_y.append(T_i)
else if Jsim(T_i, GO) < 0.25
A_y.append(T_i)
Step 8. Distractors generation based on S_y and A_y sets.
end

The words with greater than 0.75 similarities are the words that are semantically at a very close distance to the actual key and are Synonyms. A threshold of 0.25 is given similarly for antonyms as these words are effectively antonyms of the key and can be used as proper distractors. The words appearing in the range of 0.25-0.75 are the words that have significantly less similarity to the key to be Synonyms. In addition to the lower similarity, these words also do not have enough lower similarity to be proper Antonyms. Hence, these words are not considered for the implementation as these words effectively hinder the performance of the system. With all these Synonyms and Antonyms updated frequently in the Ontology, effective distractors are formed every time and the usage of an ever-updating ontology makes the entire process dynamic and yields significantly higher accuracy compared to existing models. The implementation provides a dynamic and cohesive approach to bridge the gap between the generation of relevant questions and prompt distractors. The ontology is updated dynamically and paves way for further implementation and can further be worked on.

V. PERFORMANCE ANALYSIS AND RESULTS

The implementation was carried out for Automobile Engineering as a domain of choice, which contains topics, sub-topics and auxiliary topics as depicted in Table 3. and the relevant questions are promptly generated with appropriate distractors. Experimentations were carried out for 1892 documents that were crawled as web corpus. The Domain Ontology comprised of 117 principal concepts with 347 sub-concepts. The Granularized ontology comprised of 2783 principle concepts which make sit quite dense. Precision, Recall, Accuracy, and F-Measure were used as prospective metrics for the proposed OntoQuest framework. Equations (1), (2), (3) and (4) depict Precision, Recall, Accuracy, and F-Measure respectively.

Table 3. Generated Sub-topics for various considered domains.

Broad Domains covered	***No. of Topics***	***No. of sub-topics***	***No. of auxiliary topics***
S.I engines	42	60	50
Vehicle Dynamics	37	57	43
Propulsion	35	63	46
Suspension	39	56	48
Turbomachines	43	58	55
C.I. engines	38	61	53
Jet Propulsion	45	60	50
I.C. engines	40	58	47

From Table 4, it is clearly inferable that the proposed onto-quest framework yields an overall precision of 95.08% with a recall of 95.65% and an accuracy of 95.37% is achieved with an average F-measure of 95.35%. The reason for a high accuracy rate is owing to the reason that Domain adherent, dense ontologies were used. Also, dynamic knowledge was supplied through the Ontology Granularization process. The key generation was carried out using Jaccard similarity, which is a group-wise similarity measure with a threshold of 0.75 for Synonymizing. Similarly, a threshold of 0.25 was set for distractors. There is a 0.50 distance deviation between that of distractor and key generation. Moreover, a term frequency driven summarizer has been employed, which yields high-performance measures. The reason for having a very high threshold for synonymizing is that the terms must have a minimum deviation when it comes to choosing the appropriate key.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (1)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (2)$$

$$Accuracy = \frac{Precision + Recall}{2} \quad (3)$$

$$F - measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4)$$

Table 4. Performance of the Proposed OntoQuest Framework

Metrics	***Summarization***	***Key generation***	***Distractor Generation***
Precision	92.34	94.04	93.75
Recall	95.76	96.87	96.34
Accuracy	94.05	95.46	95.05
F-measure	94.01	95.44	95..3

The entire calculation of results is done keeping in mind, the salient features of questions, which include the relevancy of the generated question with respect to the given text corpus, and that the question formulated is informative enough to provide the necessary hint to the learner to decipher it and approach the possible answer choices properly. The reason for a high percentage of performance measures of the OntoQuest is that it is an ontology focussed method that dynamically generates granular ontology from domain ontologies using WordNet. Also, a semantic similarity approach is followed for integrating ontologies. The questions are quite distinctively generated using various keywords based on synonymizing of terms in the first level for key generation and adopting antonymization for distractor generation. However, varied thresholds are also followed to enhance the quality of the correct answers and reduce the deviation of wrong answers. Moreover, the anonymization generates three distractors which are quite similar and confuses the users. Owing to the incorporation of Ontology Granularization and usage of both static and dynamic knowledge, the proposed OntoQuest has a much higher Precision, Recall, F-Measure and Accuracy values.

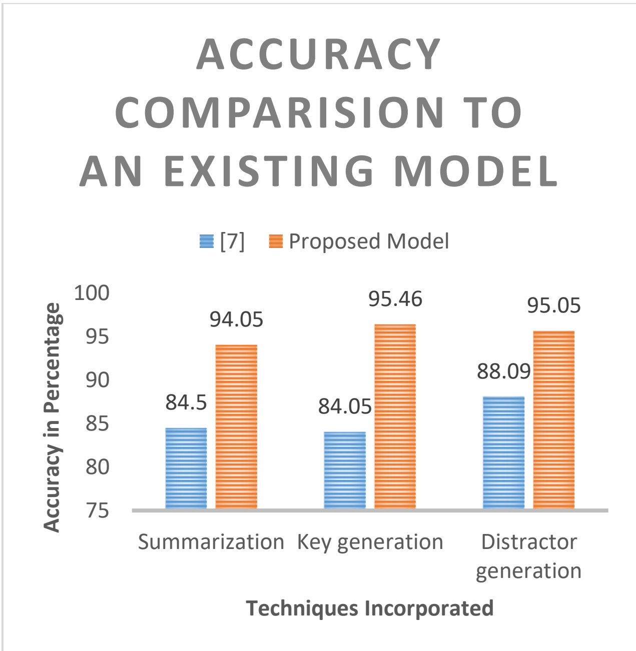

Figure 2: Accuracy comparison with an existing Model

Figure 2 depicts the comparison of the proposed OntoQuest framework with that of the FuzzyMCS approach hybridized with a cuckoo search [7]. Although the FuzzyMCS with hybridized cuckoo search incorporates ontologies, only a very basic stint of static domain ontologies has been used. This is one of the reasons why FuzzyMCS and cuckoo search is necessary for the baseline approach. However, the incorporation of a static domain ontology and increasing its density by populating term features using a dynamic lexical knowledge base and enhancing it by using a trained blooms taxonomy template increases the performance of the proposed OntoQuest. Moreover, since FuzzyMCS and other optimization algorithms are imbibed into the approach, OntoQuest is computationally less expensive. Moreover, Ontology Granularization increases the depth of the ontology and fine-tunes the amount of knowledge available. Most importantly, the summarization algorithm based on set partitioning on the basis of term frequency, makes OntoQuest quite simple and light. This makes the entire model dynamic which enhances the accuracy obtained.

VI. CONCLUSIONS AND FUTURE WORKS

An innovative framework for automatic generation of Multiple-Choice Questions for e-assessment from web corpus has been proposed. A novel strategic summarization algorithm based on POS-tagged equivalence set partitioning and term frequency has also been proposed. The other novel contributions include the use of Domain Ontologies, Ontology Granularization and Semantic Similarity based key and distractor generation techniques. The supplication of auxiliary dynamic knowledge in the premise of static ontological knowledge is an added novelty in the framework. OntoQuest is definitely a reliable model with an overall accuracy of 95.46 % and 95.05% for key and distractor generation respectively which is higher and consistent when compared to the existing models available.

ACKNOWLEDGEMENT

I thank God the Almighty and Eternal Father and my Lord Jesus Christ for giving me the strength, health, wisdom and grace to complete this paper with utmost ease and perfection. I also thank my parents and sister for standing by me and support me in all endeavours.

REFERENCES

[1] Pan, B., Li, H., Yao, Z., Cai, D. and Sun, H., 2019. Reinforced Dynamic Reasoning for Conversational Question Generation. arXiv preprint arXiv:1907.12667.

[2] Beason, William A., Swaminathan Chandrasekaran, Anne E. Gattiker, Lakshminarayanan Krishnamurthy, and Sridhar Sudarsan. "Domain-specific question-answer pair generation." U.S. Patent Application 10/170,014, filed January 1, 2019.

[3] Santhanavijayan, A., and S. R. Balasundaram. "Multi Swarm Optimization based Automatic Ontology for E-Assessment." Computer Networks (2019).

[4] Diatta, Baboucar, Adrien Basse, and Samuel Ouya. "Bilingual Ontology-Based Automatic Question Generation." In 2019 IEEE Global Engineering Education Conference (EDUCON), pp. 679-684. IEEE, 2019.

[5] Leo, Jared, Ghader Kurdi, Nicolas Matentzoglu, Bijan Parsia, Ulrike Sattler, Sophie Forge, G. Donato, and Will Dowling. "Ontology-Based Generation of Medical, Multi-term MCQs." International Journal of Artificial Intelligence in Education 29, no. 2 (2019): 145-188.

[6] Alberti, Chris, Daniel Andor, Emily Pitler, Jacob Devlin, and Michael Collins. "Synthetic QA Corpora Generation with Roundtrip Consistency." arXiv preprint arXiv:1906.05416(2019).

[7] Santhanavijayan, A., and S. R. Balasundaram. "Fuzzy-MCS algorithm-based ontology generation for e-assessment." International Journal of Business Intelligence and Data Mining14, no. 4 (2019): 458-472.

[8] Deepak, G., Shwetha, B.N., Pushpa, C.N., Thriveni, J. and Venugopal, K.R., 2018. A hybridized semantic trust-based framework for personalized web page recommendation. International Journal of Computers and Applications, pp.1-11.

[9] Pushpa, C. N., Gerard Deepak, J. Thriveni, and K. R. Venugopal. "Onto Collab: Strategic review oriented collaborative knowledge modeling using ontologies." (2015): 1-7.

[10] Deepak, Gerard, J. Sheeba Priyadarshini, and MS Hareesh Babu. "A differential semantic algorithm for query relevant web page recommendation." In 2016 IEEE International Conference on Advances in Computer Applications (ICACA), pp. 44-49. IEEE, 2016.

[11] Gulzar, Zameer, A. Anny Leema, and Gerard Deepak. "Pcrs: Personalized course recommender system based on hybrid approach." Procedia Computer Science 125 (2018): 518-524.

[12] Deepak, Gerard, and J. Sheeba Priyadarshini. "Personalized and Enhanced Hybridized Semantic Algorithm for web image retrieval incorporating ontology classification, strategic query expansion, and content-based analysis." Computers & Electrical Engineering 72 (2018): 14-25.

[13] Deepak, G., Ahmed, A. and Skanda, B., 2019. An intelligent inventive system for personalised webpage recommendation based on ontology semantics. International Journal of Intelligent Systems Technologies and Applications (IJISTA), 18(1/2), pp.115-132

[14] Giri, G.L., Deepak, G., Manjula, S.H. and Venugopal, K.R., 2018. OntoYield: A Semantic Approach for Context-Based Ontology Recommendation Based on Structure Preservation. ICCIDE 2017, p.265.

[15] Deepak, G. for user targeted e-commerce. International Journal of Computer Aided Engineering and Technology, 11(4/5), pp.449-466.

[16] Pushpa, C.N., Deepak, G., Thriveni, J. and Venugopal, K.R., 2016. A Hybridized Framework for Ontology Modeling incorporating Latent Semantic Analysis and Content based Filtering. International Journal of Computer Applications, 150(11).

[17] Deepak, G. and Gulzar, Z., 2017. Ontoepds: Enhanced and personalized differential semantic algorithm incorporating ontology driven query enrichment. Journal of Advanced Research in Dynamical and Control Systems, 9(Specia), pp.567-582.

[18] Deepak, G. and Priyadarshini, S., A hybrid framework for social tag recommendation using context driven social information