www.ijaceeonline.com ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 4, Issue 2, February 2019, pp.1–5

Text Summarization using Machine Learning Approaches for Question Answering System

K. Karpagam
Assistant Professor, Department of MCA,
Dr. Mahalingam College of Engineering and Technology, Pollachi, India
Email: karpagam80@gmail.com

Dr. A. Saradha
Professor and Head, Department of Computer Science and Engineering,
IRTT, Erode, India
Email: Saradha.irtt@gmail.com

Abstract: *Recent days, most of the user queries are of complex in nature due to increase in need of dynamic data. The complex questions are answered with list of relevant sentences from the internet documents. These can be answered by community question answering (CQA) such as Quora, stack overflow, Yahoo! Answers etc for various user levels using machine learning techniques. Answers can be generated from various sources and summarized from list of sentences. This paper proposes a novel complex question answering system which focuses on automatic machine-generated summaries. Text summarization is the technique used for summarizing text with unique, relevant features in compression ranges. The learning model is trained with benchmark datasets 20newsgroup and DUC2001 using machine learning algorithms. The experiments are carried out and verified with standard metrics such as ROUGE for the outcomes.*

Keyword: *Information Retrieval; Text Summarization; Machine Learning; Cross- Fold validation; ROUGE*

1. INTRODUCTION

The needs of web resources are growing enormously, since the large user of origin needs to obtain information with accuracy and reliable resources. But the real scenario is that users were lost while looking for information in a large group of data. Question answering system supports the understanding of the natural language of the user's query and display precise answers. It is categorized into open-Domain QA which the queries can to related to any field and answers from unstructured document or web sources and Closed Domain QA systems which are able to answer questions related to specific domain like Commercial, education, music, weather forecasting, Tourism, Medical health etc.

Question answering system are in greater demand, as they serve several interdisciplinary domains to provide support from academics to corporate for exacting precise information. To generate answers, machine learning is used for easy retrieval with an appropriate training data set and is evaluated with a validation set.

Machine learning models can be carried out by following methods such as supervised learning, unsupervised learning and reinforcement learning.

The supervised learning model works as input and output data from the system are known in advance, for example, they are used in clustering applications and predictive analysis. The unsupervised learning input data is known and the output data is unknown, for example, it is used in descriptive analysis. Reinforcement learning is the highest level of learning, both the range of data entry and the output data are unknown. For example, it is used in chess games in which the movement of the system is based on the movement of the user. It is done by self-learning the machine whenever necessary, which will increase the productivity of the system.

Complex user queries because cannot be answer with a single sentence so need of summarization arises to combine data from various resources to answer generation. For example, Why sky is blue in nature? cannot be answer in a single sentence, it need a descriptive answer. The process of text summarization is extracting the contents from original text and grouping in a shorter form with compression ratio ranges. Text summarization technique provides useful information for complex questions.

In two types, the text summary is made as extractive and abstractive. The extraction method that takes related sentences from the input document and generates a list of unique sentences. An abstractive method that provides a text summary based on original text in

Cite this paper:

K. Karpagam, A. Saradha, "Text Summarization using Machine Learning Approaches for Question Answering System", International Journal of Advances in Computer and Electronics Engineering, Vol.4, No. 2, pp. 1-5, February 2019.

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 4, Issue 2, February 2019, pp.1–5

the form of reorganized words. For large corpora, the methods are incorporated with an optimization technique to return sentences with the best scores.

The Reinforcement learning is one of the best machine learning methods for summarization with reward functions. In which uses algorithm to develop optimal solutions through actions. It is widely used in games like chess, balancing a pole, product comparison in online shopping, text and web mining. The difference between supervised learning and reinforcement learning is that reinforcement uses reward function which acts as a feedback to the candidate answer analysis. While supervised learning provides output only without feedback for the inputted text.

2. RELATED WORKS

In Paper [1], the author adopted the extractive multi-document summary techniques and formulated reinforcement learning problem to answer complex questions. Techniques for selecting an appropriate sentence using a modified linear, gradient-descending version of the Watkins Q (λ) algorithm along with a greedy strategy to determine the best possible action state. Performance is compared to Support Vector Machine, rated and improved by ROUGE score.

In Paper [2], the author proposed an approach that addresses the problem of multi-document summarization by developing a framework with user request, corpus, domain-independent techniques, statistical processing. The metrics to reduce redundancy among fetch sets and maximize diversity in the selected passages.

In paper [3] and [4], constructs a new approach for automatic text summarization using Reinforcement Learning, which optimize the candidate answer summary with score and reward function. Sentence compression is carried out with uniqueness and answer size up to 250 words. The performance is measured with DUC 2004 and compared with ROUGE scores.

In paper [5], propose a multiple-document summarization system with user interaction for choosing the keywords on the screen. For selected keywords, the user requirement is clearly understood and produced answers summaries. The 12 keyword limit is fixed and performance of system is evaluated by scoring function.

The goal of this paper [6] is to compare human generated and machine generated summaries by using different automatic text summarization evaluation techniques like ROUGE is applied. The human generated summaries are obtained from English teachers and automatic summaries are obtained using Fuzzy method & Vector approach.

In paper [7], the authors propose the generic document summarization and extractive summarization method which is based on continue sentence clustering. In paper [13] and [14] discusses on reinforcement learning model which uses complex questions and the corresponding human-generated summaries for training. The machine generated summaries generated by are evaluated for efficiency by reward function and feature weights.

3. SYSTEM ARCHITECTURE

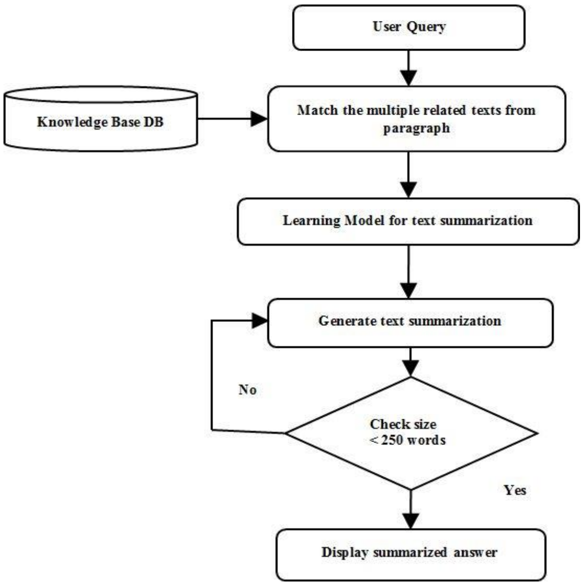

Figure 1: Proposed Architecture

The process of the proposed system is shown in Figure 1. User query is taken as natural language input, preprocess the query, extract keywords and question identification. Knowledgebase is developed using benchmark datasets. The answer is extracted from knowledge base, in which turn goes to extract the sentences based on semantic and syntactic nature. Knowledgebase is formed in cluster methods which help in easy retrieval. The distances between the similar clusters are calculated by Euclidean distance. It is calculated by the formula (1)

$$d_m\left(\mathbf{x}^r,\mathbf{x}^s\right)=\left[\sum_{j=1}^{d}\left(x_j^r-x_j^s\right)^p\right]^{1/p} \quad \text{\_\_\_\_(1)}$$

Where the X^r, X^s are the two cluster instances, p is the probability of occurrences; m is the total number of clusters. The extractions of sentences are done by applying the machine learning by developing the learning model. In complex situations the humans can perform better than machines and complex question needs deep analysis for efficient results. To achieve this, train the learning model with human-generated answers with query as input. To mitigate this, combine the related answer sentences together which should be non-redundant, relevant, informative and unique relevant sentences are arranged in chronological order.

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 4, Issue 2, February 2019, pp.1–5

Learning model is trained with the 500 questions & sample generated human summaries and tested accordingly. The sentence selection has the features of word order, length and position. The sentences count is stop when it reaches 250 words of data. with compression ratio ranges from 5% to 30%. While compressing the text, the input given is document to be summarized with ratio or number of sentences to be produced. The accuracy of machine generated summaries is compared with the human generated summaries on range of ratios. The answer is displayed to the user as useful information.

4. EXPERIMENTAL RESULTS AND PERFORMANCE EVALUATION

In papers [11] & [12] discusses on various evaluation metrics such as Precision, Recall and F1 for verifying the accuracy of the answers. Precision calculates the true positive divided by the sum of false positive and true positive. It is calculated using the Equation (2)

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

Recall is calculated as the number of correct positive predictions divided by the total number of true positives and false negatives. It is calculated using the Equation (3).

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

The F1-Score is the combination of precision and recall into a single score by calculating different types of means of both metrics. It is calculated by Equation (4)

$$\text{F1-score} = 2 * \frac{precision * recall}{precision + recall} \tag{4}$$

TABLE I COMPARING GENERATED SUMMARIES

Attribute	Source text	Human summary	Machine summary
Number of sentences	10	4	3
Number of words	250	65	57
Precision	0.38		
Recall	0.32		
F1- score	0.23		

The Table I shows the comparison between the human generated summary and machine generated summary from the single source document text. The evaluation with the metrics for checks its accuracy also examined.

The accuracy is calculated with standard metrics like the precision, F1-score and recall values with rates ranges from 5% to 30% for 10 summaries are shown in the Table II.

TABLE II COMPARESSION RATIO ACCURACY RESULTS

	Precision	Recall	F1-score
Compression rate:5%	0.7	0.6363	0.6666
Compression rate:10%	0.4	0.3636	0.3809
Compression rate:20%	0.9	0.6	0.72
Compression rate:30%	0.6	0.4285	0.5

4.1 Dataset

The knowledge base is generated using benchmark datasets used in 20newsgroups and DUC2001.The 20newsgroups data set from the UCI repository, which consists of different types of domain information such as politics, entertainment, business, technology and sports, has been used as a source of documents. The test is applied to the 20newsgroup benchmark data set and passed to the abstracts. The 20newsgroup dataset raw documents were extracted into 5 domains like sports, entertainment, politics etc for easy retrieval of data [8]. The text summarization is evaluated with DUC2001 from AQUAINT corpus which is comprised of 45 topics with new articles (each topic with 25 documents). The human summaries are generated and evaluated with the machine generated summaries with quality and responsiveness [9].

4.2 Manual evaluation

The human generated summaries are evaluated with the standards defined by NIST such as grammaticality, non-redundancy, referential clarity, focus, structure, and coherence. Judges will allot a score based on their content accuracy and relevance. The proposed system is evaluated with a pretest with the group of 10 learners who use the system for cracking competitive exams. The test conducted with 100 queries as input to the system on various domains and analysis their accuracy of summarization.

4.3 Automatic evaluation

The ROUGE (Recall-Oriented Understudy for Gisting Evaluation), a software package introduced by Lin in 2004 automatically determines the quality of text summarization in natural language [10]. This ROUGE can be evaluated with ROUGE-N (N-gram based co-occurrence statistics), ROUGE-L (Longest Common sentence Subsequence), and ROUGE-S (Skip-bigram-based co-occurrence statistics).The pro-

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 4, Issue 2, February 2019, pp.1–5

posed system used ROUGE-L (sentence level of longest common sequence) for estimated with F-measure to find the similarity between two summaries X of length m and Y of length n, where X is a human summary sentence and Y is a machine summary sentence. It is calculated by the empirical formula as follows:

$$R_{lcs} = \frac{LCS(X/Y)}{m} \tag{5}$$

$$P_{lcs} = \frac{LCS(X,Y)}{n} \tag{6}$$

$$F_{lcs} = \frac{(1+\beta^2)R_{lcs}P_{lcs}}{R_{lcs}+\beta^2 P_{lcs}} \tag{7}$$

Where LCS(X, Y) denotes the length of longest common subsequence of X and Y.

The performance evaluation of the proposed system Reinforcement Learning with Context Summarization (RLCS) is compared with the existing system accuracies such as baseline, Support Vector Machine (SVM), Reinforcement Learning (RL). Results are shown in following Table III.

TABLE III: PERFORMANCE COMPARISON: F- MEASURE

Systems	ROUGE-L
Baseline	0.21651
Support Vector Machine(SVM)	0.34313
Reinforcement Learning (RL)	0.37347
Proposed model (RL + Context Summarization(RLCS))	0.39389

5. CONCLUSION

This paper contributes novelty to the text summarization through Reinforcement Learning with Context Summarization. Reinforcement learning is used to summarize the source text using a learning model. Learning model is developed using a training set of data with complex questions and their corresponding summaries. Bench mark datasets are used for the development of the knowledge base generating answer. Complex question can be answered by summarizing the list of unique, relevant sentence. The complex question can be answered by summarizing the list of unique and relevant sentences. This proposed system has a compression range of up to 250 words or 30% of the original text. The results of the experiment show that the external system performs a summary of the text with automatic learning techniques.

REFERENCES

[1] Chali, Y., Hasan, S. A., & Imam, K,"A reinforcement learning framework for answering complex questions", *In Proceedings of the 16th international conference on intelligent user interfaces*, IUI '11, ACM 2011, pp. 307–310.

[2] Goldstein, J., Mittal, V. Carbonell, J., & Kantrowitz. M.," Multi-document summarization by sentence extraction", Proceedings of the 2000 NAACLANLP Workshop on automatic summarization, Vol. 4, ACL, 2000, pp. 40–48..

[3] Ryang, S., & Abekawa, T., "Framework of automatic text summarization using reinforcement learning", *In Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning, EMNLP- CoNLL* Jeju Island, Korea ,2012, pp. 256 265.

[4] Cody Rioux et .al ," Fear the REAPER: A System for Automatic Multi-Document summarization with Reinforcement Learning" , *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, ,October 25-29, 2014, PP 681–690.

[5] Sakai, H., & Masuyama, S,"A Multiple-document summarization system with user interaction", *In Proceedings of the 20th international conference on computational linguistics*,COLING '04, ACL, 2004.

[6] Farshad Kiyoumarsi," Evaluation of Automatic Text Summarizations Based on Human Summaries", 2*nd GLOBAL CONFERENCE on LINGUISTICS and FOREIGN LANGUAGE TEACHING, LINELT-2014, Procedia - Socia and Behavioral Sciences*, ELSEVIER 2014, PP. No 83 – 91.

[7] R.M.Aliguliyev,"A new sentence similarity measure and sentence based extractive technique for automatic text summarization", *Expert System Applications* 2009, 36(4),PP 7764-7772.

[8] http://qwone.com/~jason/20Newsgroups/

[9] https://github.com/tapilab/is-karthikbmk

[10] Lin, C," ROUGE: A package for automatic evaluation of summaries", In *Proceedings of workshop on text summarization branches out, post-conference workshop of association for computational linguistics*, Barcelona, Spain 2004, pp. 74–81.

[11] lomiyets, Oleksandr & Moens, Marie-Francine,"A survey on question answering technology from an information retrieval perspective", Information Sciences, 2011, Volume 181, Issue 24, 15 December 2011, Pages 5412-5434.

[12] Tilani Gunawardena, "Performance Evaluation Techniques for an Automatic Question Answering System", *International Journal of Machine Learning and Computing*, Vol. 5, No. 4, August, 2015.

[13] Yllias Chali et al,"A reinforcement learning formulation to the complex question answering problem", *Information Processing and Management, Elsevier,* Volume 51, Issue 3, May 2015, Pages 252-272.

[14] Chali, Y., Hasan, S. A., & Imam, K.," Improving the performance of the reinforcement learning model for answering complex questions", I*n Proceedings of the 21st ACM conference on information and knowledge management,* CIKM 2012, ACM, 2012,pp. 2499– 2502.

Authors' Biography

K. Karpagam is working as Asst. Professor in Department of Computer Applications, Dr. Mahalingam College of Engineering and Technology, Pollachi. She has 15 years of academic experience and currently pursuing research in Anna University, Chennai. The area of research interests includes text mining, Question Answering system and Human computer Interaction.

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 4, Issue 2, February 2019, pp.1–5

Dr. A. Saradha is working as Professor and HOD, Department of Computer Science and Engineering, IRTT, Erode. She has 25+ years of academic experience. Her research interests are Semantic Web, Human Computer Interaction, Image processing.

Cite this paper:

K. Karpagam, A. Saradha, "Text Summarization using Machine Learning Approaches for Question Answering System", International Journal of Advances in Computer and Electronics Engineering, Vol.4, No. 2, pp. 1-5, February 2019.