

Analysis of Comprehension Solutions for the Pali Context of Karaniya Metta Sutta

Anupama Sakhare¹

Assistant Professor Department of Computer Science
R.T.M. Nagpur University, Nagpur

Abstract

Background: The present work is about reading comprehension task. In this system the task of answering questions is performed which is based on an inputted passage. This system understands the given passage. In the present work the passage is the text of Karaniya Metta Sutta from the corpus of 84,000 Teachings of Buddha.

Materials and Methods: The first verse is chosen as the text for performing the Question-Answer system passage. Three methods are used in the study :

- 1) ELMo method
- 2) BiDAF Method
- 3) NAQANet method

The experiments are performed using Allen NLP Demo version. The tool used for the study is "Reading Comprehension" system.

Result: The analysis reveals that when the same input of first verse of Karaniya Metta Sutta is inputted the maximum accuracy.

Conclusion: The accuracy of 99% is obtained using the ELMo method followed by that BiDAF method. However NAQANet does not produce the expected result.

Keywords: Passage context, Karaniya Metta Sutta, Question-Answering System.

Date of Submission: 18-05-2020

Date of Acceptance: 03-06-2020

I. Introduction

A Natural language processing system is about- Understanding a Human Language. It is used for morphological Analysis, Syntactic Analysis, Discourse Integration and Pragmatic Analysis.

Question-Answering system is one of the applications of the Natural language understanding system. The passage-context is the text which is inputted in the system. This input is preceded by a question which is based on Reading Comprehension. After the Question for Answering it a set of model is available. One of the method is to be chosen for performing the task of Question-Answering.

In the present study the following passage context is selected. The passage inputted into the system is "He who is skilled in doing good and who wishes to attain that state of calm should act thus- He should be able, upright, obedient, gentle and humble."

This is the Roman version (Translation) of the first verse/section of the Karaniya Metta Sutta.

II. Method And Models

In this method the Question also needs to be inputted.

The Question is

"How should one act?"

Followed by the Question the first model is chosen.

ELMo is a deep contextualized word representation that models both (1) complex characteristics of word use (e.g., syntax and semantics), and (2) how these uses vary across linguistic contexts (i.e., to model polysemy). These word vectors are learned functions of the internal states of a deep bidirectional language model (biLM), which is pre-trained on a large text corpus. They can be easily added to existing models and significantly improve the state of the art across a broad range of challenging NLP problems, including question answering, textual entailment and sentiment analysis.

Machine comprehension (MC), answering a query about a given context paragraph, requires modeling complex interactions between the context and the query. Recently, attention mechanisms have been successfully extended to MC. Typically these methods use attention to focus on a small portion of the context and summarize it with a fixed-size vector, couple attentions temporally, and/or often form a uni-directional attention. In this

paper we introduce the Bi-Directional Attention Flow (BiDAF) network, a multi-stage hierarchical process that represents the context at different levels of granularity and uses bi-directional attention flow mechanism to obtain a query-aware context representation without early summarization. Our experimental evaluations show that our model achieves the state-of-the-art results in Stanford Question Answering Dataset (SQuAD) and CNN/DailyMail cloze test.

With system performance on existing reading comprehension benchmarks nearing or surpassing human performance, we need a new, hard dataset that improves systems' capabilities to actually *read* paragraphs of text. DROP is a crowd-sourced, adversarially-created, 96k-question benchmark, in which a system must resolve references in a question, perhaps to multiple input positions, and perform discrete operations over them (such as addition, counting, or sorting). These operations require a much more comprehensive understanding of the content of paragraphs than what was necessary for prior datasets. AllenNLP provides an easy way for you to get started with this dataset, with a dataset reader that can be used with any model you design, and a reference implementation of the NAQANet model that was introduced in the DROP paper.

(1) ELMo model :

When the ELMo model is selected the output is generated in the form of Answer which is –
“act thus – He should be able, upright, perfectly upright, obedient, gentle and humble.”

(2) BiDAF model (trained on SQuAD) :

For the same passage context when BiDAF model is selected the following Answer is obtained :
“able, upright, perfectly upright, obedient, gentle and humble”

(3) NAQANet (trained on DROP) :

The first section of Karaniya Metta Sutta is inputted as the passage context and the Question inputted is : “how should one act.” The model selected is NAQANet model. For this Question the model answers :
“State of calm should act” which is not the perfect answer.

III. Result

The same passage context was inputted and 3 different methods were chosen, one at a time to give the correct Answer result. It is found that the first two methods viz. (1) ELMo and (2) BiDAF give better result as a reading comprehension methods as compared to the third method (3) NAQANet method.

IV. Conclusion

Question –Answering system is a very important application and it can be used in various computer-based tutor packages for evaluating the correct answers.

Acknowledgement

The author is thankful to the developers of Allen NLP Lab.

References

- [1]. <https://demo-allennlp.org/reading-comprehension>
- [2]. <https://www.nltk.org/>
- [3]. <https://towardsdatascience.com/gentle-start-to-natural-language-processing-using-python-6e46c07addf3>
- [4]. <https://www.lexalytics.com/lexablog/machine-learning-natural-language-processing>
- [5]. Kamal Nigam, et al. : Text Classification from Labeled and Unlabeled Documents using EM. Machine Learning, 1999
- [6]. Peter F Brown, et al. : Class-Based n-gram Models of Natural Language, 1992
- [7]. Marti A. Hearst : Automatic Acquisition of Hyponyms from Large Text Corpora, COLING, 1992
- [8]. Peter F Brown et al. : A Statistical Approach to Machine Translation, Computational Linguistics, 1990

Anupama Sakhare. “Analysis of Comprehension Solutions for the Pali Context of Karaniya Metta Sutta.” *IOSR Journal of Computer Engineering (IOSR-JCE)*, 22(3), 2020, pp. 36-37.