

Paraphrased Question Answering Using Recurrent Neural Network in Bangla Language

Submitted By

Nazmus Sakib Patwary

ID: 2015-2-60-092

Md. Mohaiminul Hasan

ID: 2015-2-60-084

Tanvir Rahman

ID: 2015-2-60-027

Supervised By

**Md. Mohsin Uddin
Lecturer**



**Department of Computer Science and Engineering
East West University
Dhaka-1212, Bangladesh**

December, 2019

Declaration

We, hereby, declare that the work presented in this thesis is the outcome of the investigation performed by us, under the supervision of Md. Mohsin Uddin, Lecturer, Department of Computer Science and engineering, East West University. We also declare that no part of this thesis has been or is being submitted elsewhere for the award of any degree or diploma.

Countersigned

.....

(Md. Mohsin Uddin)

Supervisor

Signature

.....

(Nazmus Sakib Patwary)

(**2015-2-60-092**)

Signature

.....

(Md. Mohaiminul Hasan)

(**2015-2-60-084**)

Signature

.....

(Tanvir Rahman)

(**2015-2-60-027**)

Letter of Acceptance

We, hereby, declare the thesis is from the student's own work and best effort of us and all other sources of information used in this paper have been acknowledged. This thesis has been submitted with our approval.

Supervisor
Signature

(Md. Mohsin Uddin)
Lecturer
Department of Computer Science and Engineering
East West University, Dhaka

Chairperson
Signature

(Dr. Taskeed Jabid)
Associate Professor
Department of Computer Science and Engineering
East West University

Abstract

Recent studies on QA (Question Answering) system in English language have been emerged extensively with the composition of NLP (Natural Language Processing) and IR (Information Retrieval) by amplifying miniature sub tasks to accomplish a whole AI-system having capability of answering and reasoning complicated and long questions through understating paragraph. In our proposed study, we present a general heuristic framework, an end-to-end model used for paraphrased question answering using single supporting line which is the initial appearance ever in Bangla language. Corpus dataset was scrapped from Bangla wiki and then questions were generated corresponding context have been used to learn the model. Translated bAbI dataset (1 supporting fact) [5][6] in Bangla language has been also incorporated with to experiment the proposed model manually. To predict appropriate answer, model is trained with question-answer pair and a supporting line. For comparing our task applying variation of basic RNN (Recurrent Neural Network): LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) different accuracy has been found. For further accomplishment, synthetic and semantic word relevance in high dimension vector space: Bangla Word2vec (word embedding system) is added to the system as sentence representation along with PE (Positioning Encoding) and which outperforms both memory network GRU and LSTM precisely.

Acknowledgments

As it is true for everyone, we have also arrived at this point of achieving a goal in our life through various interactions with and help from other people. However, written words are often elusive and harbor diverse interpretations even in one's mother language. Therefore, we would not like to make efforts to find best words to express our thankfulness other than simply listing those people who have contributed to this thesis itself in an essential way. This work was carried out in the Department of Computer Science and Engineering at East West University, Bangladesh.

First of all, we would like to express our deepest gratitude to the almighty for His blessings on us. Next, our special thanks go to our supervisor, "Md. Mohsin Uddin", who gave us this opportunity, initiated us into the field of "Paraphrased Question Answering Using Recurrent Neural Network in Bangla Language", and without whom this work would not have been possible. His encouragements, visionaries and thoughtful comments and suggestions, unforgettable support at every stage of our BS.c study were simply appreciating and essential. His ability to muddle us enough to finally answer our own question correctly is something valuable what We have learned and We would try to emulate, if ever We get the opportunity.

Last but not the least, we would like to thank our parents for their unending support, encouragement and prayers.

There are numerous other people too who have shown us their constant support and friendship in various ways, directly or indirectly related to our academic life. We will remember them in our heart and hope to find a more appropriate place to acknowledge them in the future.

Nazmus Sakib Patwary

2015-2-60-092

Md. Mohaiminul Hasan

2015-2-60-084

Tanvir Rahman

2015-2-60-027

Table of Contents

Declaration of Authorship.....	i
Abstract.....	ii
Acknowledgment.....	iii
Chapter 1 Introduction.....	1-2
1.1 Overview.....	1
1.2 Objectives with specific Aims	2
1.3 Summary of Contribution.....	2
1.4 Thesis Organization.....	2
Chapter 2 Related Work.....	3-5
2.1 Introduction	3
2.2 Question Answering.....	3
2.3 Existing Works of Question Answering	4
Chapter 3 Methodology.....	6-15
3.1 Datasets.....	7
3.2 Data Preprocessing.....	8
3.3 Implementations.....	9
3.4 Predicting Answer	14
Chapter 4 Experiment and Result Analysis.....	16-18
5.1 Experiments	16
5.2 Result & Analysis	17
Chapter 5 Conclusion and Future Work.....	19
Bibliography.....	20
Appendix	22

List of Figures

1. Flow Chart of Proposed Architecture	6
2. Process diagram of positional encoding	10
3. Diagram of BOW and Skip-gram.....	11
4. Example-1 for predict result	14
5. Example-2 for predict result	15
6. Comparison view of Average Accuracy	17
7. Graphical delegation of Average Accuracy of model using LSTM and Word2Vec with 400 Epochs.....	18
8. Graphical delegation of Average Accuracy of model using GRU and Word2Vec with 400 Epochs.....	18

List of Tables

3.1 Overview of Datasets.....	08
3.2 Example of Tokenization.....	09

List of Algorithms

3.3 Gated Recurrent Unit (GRU).....	13
3.3 Long Short-Term Memory (LSTM).....	13

1.1 Overview

Intelligent question answering system that offers tasks like asking automated machine any questions and getting appropriate answer from computer automatically has been very important end-user task in recent age which ease human's life drastically. To enable communication with computer, asking question to computer is indispensable task. Many researches have been done over this particular field of question answering in near years. However, it's an important task which uses a combination of both NLP and IR that shortens the distance between IR-based search and intelligent assistants that uses information extracting process from data [3].

Different answers can be yielded with a modicum variation in semantically equivalent questions. For example, questions like "who created Microsoft" and "who started Microsoft" yield same identify. The question answering model ought to acknowledge the answer from its knowledge base considering both questions semantically equivalent [2].

There has been a lot of research in machine learning that are intended to reasoning and intelligently answering questions. It's a comprehensive area [5]. Towards answering question and reasoning two grand challenges in intelligence system have been arrived in numerous research that make models which is able to make multiple computational steps for question answering and to make model that adopts and working ability considering long term dependencies in case of sequential data as well as unstructured data. [6]. In the circumstance of semantic parsing for answering question in recent time, researchers are highly focused on complicated and long question answering. [4]

1.2 Objectives with specific Aims

In our study we propose a system having consideration of close domain dataset and develop algorithms with variety for understanding language and paraphrased question answering. Different architecture of RNN like LSTM, GRU has been used to build satisfactory model. Question-answer pair from Bangla close domain dataset has been used as training data as well as with one support line and multiple related line. It can answer paraphrased question's answer too and questions containing 'who' ('কে'), 'where' ('কোথায়'), 'when' ('কখন'), 'what' ('কি') are answered also whereas [5][6] considered so straight forward simple dataset that only contain questions containing 'where' ('কোথায়') for their single supporting fact category.

1.3 Summary of Contribution

Abundant number of researches has been performed considering English language related to different question answering task like search-based QA, factoid QA etc. in order to accomplish AI-complete question answering in isolated ways using end-to-end neural network. But in Bangla language no such research and task have been conducted regarding question answering which would contribute AI-complete question answering in Bangla language.

1.4 Outline

The remaining of the thesis is organized as follows:

Chapter 1 represents the motive of the system and also represent objectives and contribution

Chapter 2 presents previous work on Question Answering and also on use of those.

Chapter 3 Presents of Methodology with dataset overview and Implementations.

Chapter 4 Presents the Result and Analysis

Chapter 5 presents the Conclusion and Limitations

Chapter 2

Related Work

2.1 Introduction

Now, the running age is the pre-eminent era of using internet as well as using automated system that helps people to easily get required answer of any interrogatories. People are massively dependent on device or machine now a days to fulfill their quest easily without having much hassle. So, having a device with an intelligence technology that is capable of answering question in native language is always blessing. So, intelligence like question answering system in Bengali language can assist people to get any information just by typing few words which build valid question to get valid answer.

2.2 Question Answering

Intelligent question answering system that offers tasks like asking automated machine any questions and getting appropriate answer from computer automatically has been very important end-user task in recent age which ease human's life drastically. To enable communication with computer, asking question to computer is indispensable task. Different answers can be yielded with a modicum variation in semantically equivalent questions. For example, questions like "who created Microsoft" and "who started Microsoft" yield same identify. The question answering model ought to acknowledge the answer from its knowledge base considering both questions semantically equivalent.

Abundant number of researches has been performed considering English language related to different question answering task like search-based QA, factoid QA etc. in order to accomplish AI-complete question answering in isolated ways using end-to-end neural

network. But in Bangla language no such research and task have been conducted regarding question answering which would contribute AI-complete question answering in Bangla language.

2.3 Existing Works of Question Answering

In this paper [1], authors brought in a Recurrent neural network model that will predict person, any things or place related to given description about any entities. They have used huge amounts of unstructured and compositional data. They proposed dependency tree recursive neural network and they extended this to question answering neural network with trans-sentential averaging (QANTA) which can combinedly learn word and phrase-level representation to identify about entities. Their model takes dependency parse tree of sentences of questions and takes corresponding answer about entity as labeled input.

Equivalent questions with slight difference may lead to different answers in case of question answering system [2]. A question answering system must interpret between those question and should give the appropriate answer from the knowledge base. In this paper, authors proposed a model that learns paraphrases for question answering and produces an estimation of probability distribution over all candidate answers.

Another paper [3] where researchers proposed an efficient neural model denoted as FastQA for question answering which outperforms existing model over very popular recent datasets named SQuAD, NewsQA and MsMARCO. They used bi-directional RNN to learn input question and context. For answer span they used beam-search with a size K which reveals the top k -answers are predicted and the highest probability span is predicted as the appropriate answer based on learning model.

For accomplishing language understanding by the use of appealing QA dependent strategies has been emerged in nearly years [5]. In this paper [5] authors have done experiment with 20

different tasks (based on supporting fact and reasoning fact) of question answering for getting suitable models applicable to detect symbolic sequence in language. Standard MemNN outperforms over the LSTM baseline and N-gram baseline in their approach applied on simple bAbI dataset. Using the same dataset, positioning encoding (PE) representation enhances performance over bag-of-word (BoW) for few particular tasks [6]. Raising hops also yields better accuracy for their given approach.

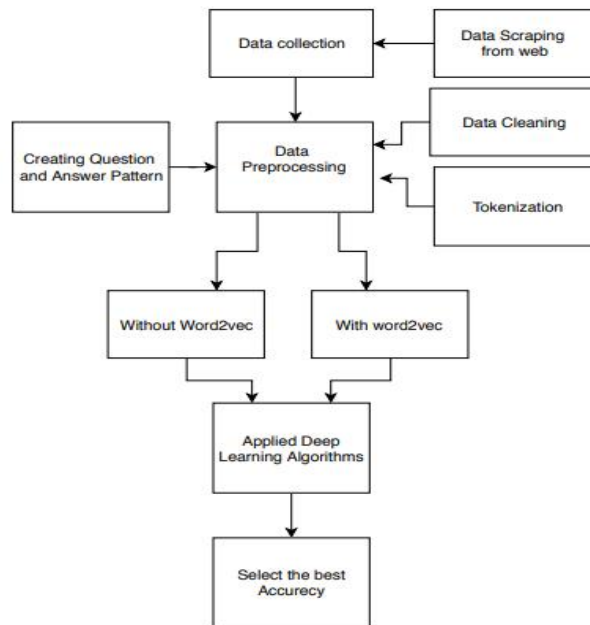


Fig. 1. Flow Chart of Proposed Architecture

In our study, several steps have been followed and the figure: -1 illustrates the whole procedure. After executing all those steps, several outcomes have been found and best few accuracies have been taken for the evaluation of all the methods. Firstly, we collected existing data set and translated into Bangla and the another data set we have created where the dataset contains many historic information. So we applied both dataset. After the data collection step, we have used several data preprocessing techniques. Then we applied two Deep Learning techniques using both position

encoding and word embedding(word2vec). In this process, we have created several combination using activation function and optimizer of both LSTM and GRU algorithms.

3.1 Datasets

We have evaluated our approach with two different datasets over Bengali Language. At very first, we have collect Bangla corpus from Bangla wiki to assessment the new approach. Based on that corpus we have generated question-answer pair as well as supporting line by hand manually for the reason no preprocessed dataset in Bangla language has been found. In our study we only consider history domain related question in Bangla language. So, our model is based on close domain QA which is prerequisite motive for a robust AI-question answering system of Bangla language. Our second dataset is bAbI dataset which is a project of towards the goal of automatic text understanding and reasoning organized by Facebook AI Research. But their available language is English and Hindi. So, we have translated the English language dataset into Bengali language with Google Translate API. This dataset is containing 20 sets of task. For our approach we have only took single supporting fact data. This dataset contains thirty thousand lines of data with ten thousand question-answer pair. Basically “Facebook bAbI” dataset is noiseless and human readable. But our corpus dataset is containing lots of noise and unnecessary symbols in data. We have purified corpus dataset removing unnecessary symbols and objects, words from other language. By doing all these cleaning stuffs our dataset got ready completely for the experimental and learning issues of the framework and it can be triggered for the further study concerning QA in Bangla language.

TABLE 1

Overview of Datasets

<i>Source of Dataset</i>	<i>Total line</i>	<i>Total words</i>	<i>Total unique words</i>	<i>Question-Answer Pair</i>
<i>History Corpus</i>	3500	30000	9878	1500
<i>Facebook bAbI Dataset</i>	30000	125000	24	10000

3.2 Data Preprocessing

In our study, we have applied two types of data, one is translated data and another one is wiki data. In dataset section, we have discussed in details about datasets. In second step, we have cleaned our data, mostly unnecessary characters or word or URL link have been removed. For each and every cases we have created pattern and simply removed the unnecessary word or characters. For the final step, we have preparing data for the training for example creating question, query pattern and tokenization.

Web scraping: For wiki dataset, firstly we scraped history of Bangladesh from wiki. Web scraping is very efficient technique to collect the data from website. It saves lots of time. Using Web scraping, can be collect lots of data from websites. In this paper we scraping history of Bangladesh from wiki.

Tokenization: Tokenization is a process by which corpus sentence is converted into meaningful pieces where each word of the corpus is considered as a token. By dividing sentences into collection of tokens following splitting by space data gets ready for the next step called word representation. For Example,

TABLE 2

Example of Tokenization

Sentence	Tokenization
I love Bangladesh.	'I' , 'love', 'Bangladesh' , '.'
আমি বাংলাদেশকে ভালোবাসি।	'আমি', 'বাংলাদেশকে', 'ভালোবাসি', '।'

3.3 Implementation

Position Encoding: Sequence to sequence model always have encoder and decoder. In sequence dependent data, the encoder has input sequence and it actually maps those data through a vector. In text data maintaining sequence of word is important for further operation. The position encoding technique actually map each word of the corpus into a vector serially so that it can remember all the position of the word in the time of generating output from the mapped value. Consider a token x_1 , which has position i , then for the next token x_2 position will be $i+1$ or j . It is different from other encoding technique. This gives model an idea about the word sequence of the corpus or sentence of corpus. Hence position notation is used to mark the position to help the learning model to consider the sequence of word. Every word has a position in a sentence. It also represents the corresponding value of the word. It's kind of indexing system. For an example:

“আমি বাংলাদেশকে ভালবাসি,
আমি বাংলাদেশকে শ্রদ্ধা করি,
বাংলাদেশকে নিয়ে আমার অনেক স্বপ্ন,
আমি আমার দেশকে ভালবাসি”

In the given corpus above, there are 4 sentences. By mapping each unique word with an indexing value, the vector is given below:

[“আমি”: 1, “বাংলাদেশকে”: 2, “ভালবাসি”: 3, “শ্রদ্ধা”: 4, “অনেক”: 5, “করি”: 6, “নিয়ে”: 7, “আমার”: 8, “স্বপ্ন”: 9, “দেশকে”: 10]

All the words have a unique value to identify the word. This value is used to do learn a language model. For all the sentence the new generated vectors are:

[1, 2, 3],

[1, 2, 4, 6],

[2, 7, 8, 5, 9]

[1, 8, 10, 3]

So, in this way words are represented in positional encoding for learning model. In LSTM, GRU based RNN positional encoding is mostly used word representation.

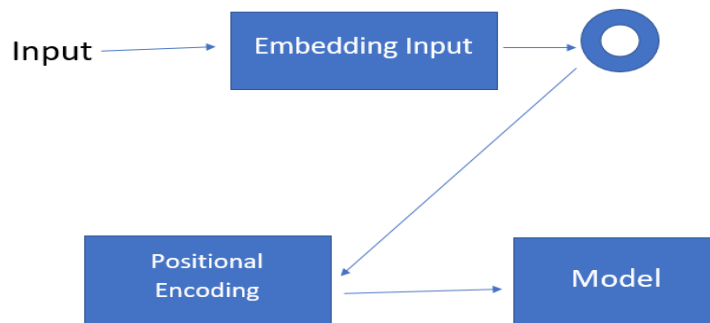


Fig. 2. Process diagram of positional encoding

Word embedding: Word embedding is also known as word2vec [12][13]. It's a cyclopean system to represent word of corpus with high dimensional vector space. It's a neural network model which contains two layers. The input of word2vec is large corpus of text and the output contains large dimensional set of vectors which is actually the feature vector of that corpus. The feature vector is then ready to be learnt through the neural language model. By using given vector cosine similarity can be measured to find the closeness between two words. Word2vec has huge impact on task like sentiment analysis and recommendation

system. The main purpose of word2vec is to cluster those words which have similar meaning based on context.

Word2vec consists of two model:

1. Skip Gram
2. Continuous bag of words (CBOW)

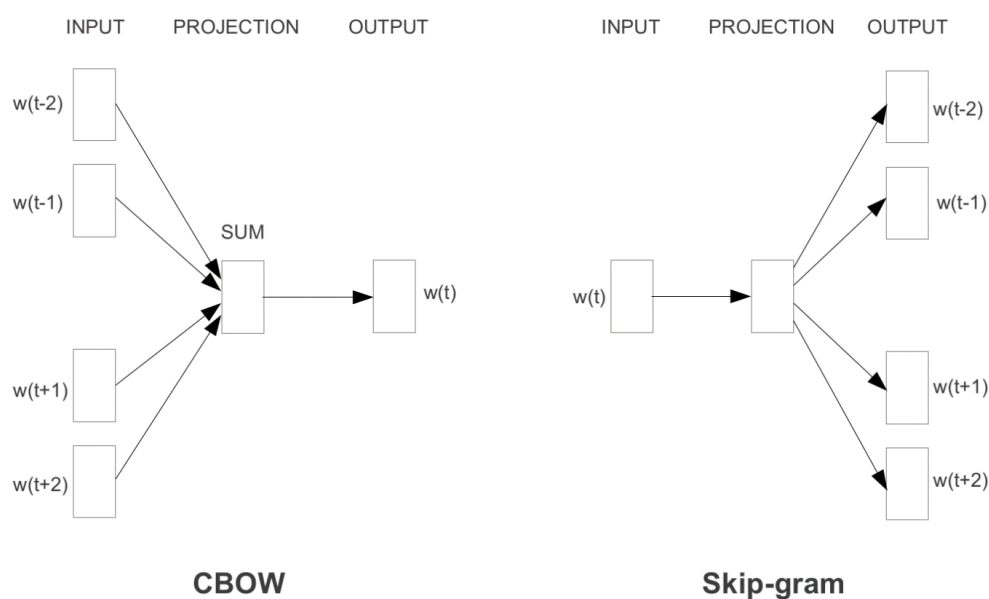


Fig. 3. Diagram of CBOW and Skip-gram

CBOW: CBOW is a model which takes bag of words as input and based on context it gives output a similar word. It tries to predict target word based on context.

Skip-gram: Skip-gram is a unsupervised model which takes a word as input and gives some words similar to that word based on given context.

In our research, Bangla word2vec has been used to embed Bangla word in 300-dimension vector space. By using this in our Bangla text corpus a precise accuracy has been got.

Bag of word: In machine learning, machine actually can't deal with string or text data. But in NLP all data are basically string or data. So, before doing learn the model, string data must be converted into numeric form. So, word to numeric conversion is one of the key terms for machine learning. To represent word, Bag of Word is one of the famous and important word representation technology of machine learning. It converts actually text corpus into vector of number. This word representation is also known as feature extraction.

Sometimes, it's called feature encoding in machine learning. So, in formal definition a BAG-OF-WORD in short BoW is a feature extraction method from text corpus in order to fit in language algorithm. BoW is consisted of two major things. One is known word's vocabulary represented in vector and another one is count of each word appears in the text corpus.

Example with a simple Bangla corpus:

“আমি”, “বাংলাদেশকে”, “ভালবাসি”, “শ্রদ্ধা”, “অনেক”, “করি”, “নিয়ে” “আমার”, “স্বপ্ন”, “দেশকে”

So, there are ten unique words in this Bangla corpus. This actually 1-gram method. Sentences are split by considering a white space here. The actual corpus contains 16 words with having same word multiple time. Now, here a vector of ten elements is considered which is given below;

[“আমি”, “বাংলাদেশকে”, “ভালবাসি”, “শ্রদ্ধা”, “অনেক”, “করি”, “নিয়ে” “আমার”, “স্বপ্ন”, “দেশকে”]

Now for all the sentences of the corpus the feature vector will be:

[1, 1, 1, 0, 0, 0, 0, 0, 0, 0],

[1, 1, 0, 1, 0, 1, 0, 0, 0, 0],

[0, 1, 0, 0, 1, 0, 1, 1, 1, 0],

[1, 0, 1, 0, 0, 0, 0, 1, 0, 1]

Now, in this way bag of word actually extract feature from text corpus. It is a feature extraction method.

Gated Recurrent Unit (GRU) [10][11]: Gated mechanism added to recurrent neural network is known as GRU. Basic recurrent network has a problem called vanishing gradient problem. To alleviate this problem GRU has been invented. It has three gates. They are:

1. Update Gate
2. Reset Gate
3. Current Memory Gate

1. **Update gate:** It is used to determine how much past memories are needed to pass into the next step of the network.
2. **Reset Gate:** It is used to determine how much past memories are needed to be forgotten.
3. **Current Memory gate:** It is used to introduce some non-linearity into the input and to also make the input Zero-mean.

Long Short-Term Memory (LSTM) [7][8][9]: The basic recurrent neural network contains a problem which is known as vanishing gradient problem. To resolve this effective problem a gated neural network has been invented by the researcher which is known as LSTM. It is consisted of:

1. Forget gate
2. Input gate
3. Cell state
4. Output gate

1. **Forget gate:** It is used to determine how much past information should be kept or deleted.
2. **Input gate:** It is used to update the cell state.
3. **Call state:** Pointwise multiplication from forget state goes to call state for the further calculation.
4. **Output gate:** The output gate decides what the next hidden state should be. Remember that the hidden state contains information on previous inputs.

3.4 Predicting Answer

Every question has a supporting number which help to find out appropriate answer. For example,

Question and answer pattern		support
1 পুঞ্জবর্ধনের রাজধানী মহাস্থানগড় ।		
2 বারো ভূঁইয়ারা কখনো মুঘলদের আনুগত্য স্বীকার করেননি এবং প্রাণ বিসর্জন দেন ।		
3 তিনি ২ হাজার অশ্বারোহী বাহিনী নিয়ে মুঘলবাহিনীর গতিরোধ করতে উষার ত্যাগ করেন ।		
4 পুঞ্জবর্ধনের রাজধানীর নাম কি?		1
Question: পুঞ্জবর্ধনের রাজধানীর নাম কি?	Answer: মহাস্থানগড়	Prediction: মহাস্থানগড়
Question: পুঞ্জবর্ধনের রাজধানী কোথায়?	Answer: মহাস্থানগড়	Prediction: মহাস্থানগড়

Fig. 4. Example-1 for predict result

In the above figure-4 has some rows and columns. In column, one is for question and another is one for supporting. In last two rows have same question but have different structure. But both different scenario, the questions gives accurate result.

Another example for the better understanding:

Question and answer pattern		support
1 মদনপাল এই বংশের শেষ রাজা ।		
2 তার পত্নী মন্ত্রীর সহযোগে বিষপ্রয়োগে স্বামী-হত্যা করেছিলেন ।		
3 ১৯৪৬ সালে উড়িষ্যা রাজ্য গঠন হয় ।		
4 উড়িষ্যা রাজ্যে কত সালে গঠন হয়?		3
Question: উড়িষ্যা রাজ্যে কত সালে গঠন হয়?	Answer: ১৯৪৬ সালে	Prediction: ১৯৪৬ সালে
Question: উড়িষ্যা রাজ্যে কবে গঠন হয়?	Answer: ১৯৪৬ সালে	Prediction: ১৯৪৬ সালে

Fig. 5. Example-2 for predict result

Chapter 4

Experiments & Results Analysis

For the experiment and analysis, we have used anaconda, which is an environment of python 3.7 with a lot of essential packages for machine learning and other useful tools.

The other configurations are:

- Processor: Intel core i5
- Clock rate: 3.4GHz
- RAM: 8GB
- Operating system: Windows (64 bit)

4.1 Experiments

First of all, for our experiment over the proposed framework, we examine two separate designs choice: (i) Positioning Encoding; (ii) Word2vec (dimension $d = 300$) for the word representation of the corpus which goes inside the model for the learning purpose in the training times.

We also contrasted two more major variations by using RNN based: (iii) LSTM; (iv) GRU - for comparing both given results from the architectural view. In LSTM, it uses three gates: input, forget, output and it uses internal memory to reminisce the past sequential effects of early layers whereas GRU uses only two gates such as update gate and reset gate which is easier to implement compare to LSTM.

By considering different activation function in the sequential model: (v) Softmax; (vi) Linear; (vii) Relu - we have got diversity of results. For separate design choice of word representation, the results of the model vary from activation function to function. For an example, for the word2vec Softmax outperforms than other activation function that reveals that Softmax deals properly with word representation in higher dimension.

4.2 Results & analysis

In this section, we compare our very first approach for the Bangla language paraphrased QA based on single support fact. We evaluate on different variation to compare the results precisely.

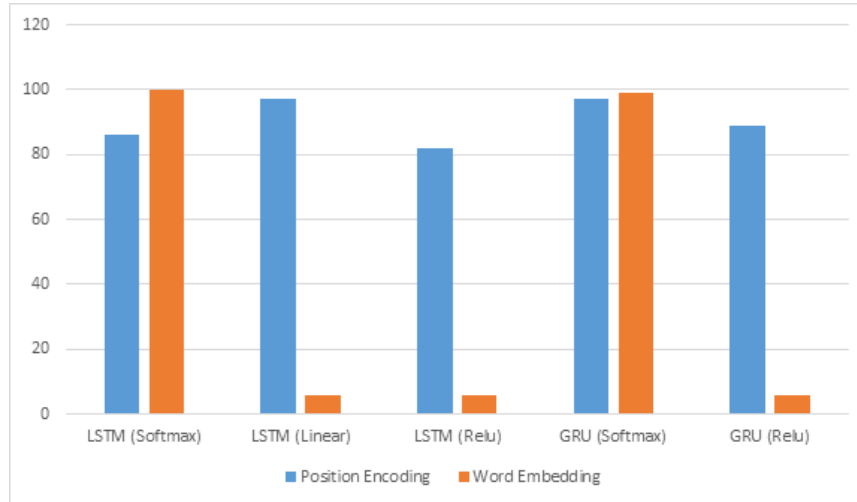


Fig. 6. Comparison view of Average Accuracy

For our dataset, Figure :-6 illustrates the overall accuracy for different approaches. For the positioning encoding representation with LSTM neural architecture we have got average accuracy for Softmax, Linear, Relu activation function respectively 86%, 97%, 82%. Using word2vec word embedding representation for same sequence we have got average accuracy 100%, 5% and 5.6%. However, Word2vec with Linear and Relu activation function causes drastically collapse in case of average accuracy. For GRU based RNN we have got average accuracy of 96.59% and 88.6% respectively for Softmax and Relu activation function in case of positioning encoding. For word embedding GRU with Softmax has given average accuracy of 98.86%. For the bAbI dataset [5,6] translated in Bangla language have gained 84% accuracy as the dataset contains only 24 unique words and word2vec doesn't fit good with that dataset. By applying their approach [5,6] in our corpus we have got 86% accuracy whereas by adding word2vec in our framework we have got 100% accuracy.

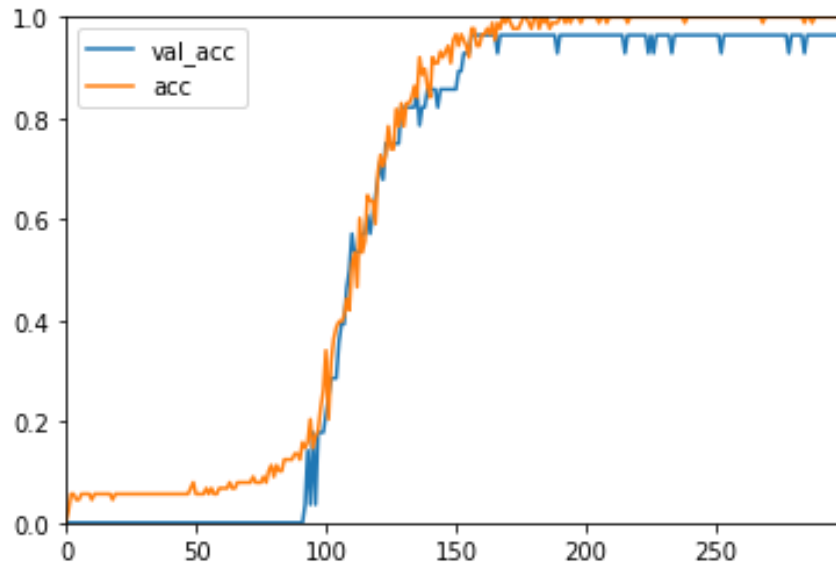


Fig. 7. Graphical delegation of Average Accuracy of model using LSTM and Word2Vec with 400 Epochs.

Figure:-7 shows the graphical representation of our model which gives the highest accuracy (100%) with the combination of LSTM and word2Vec word embedding.

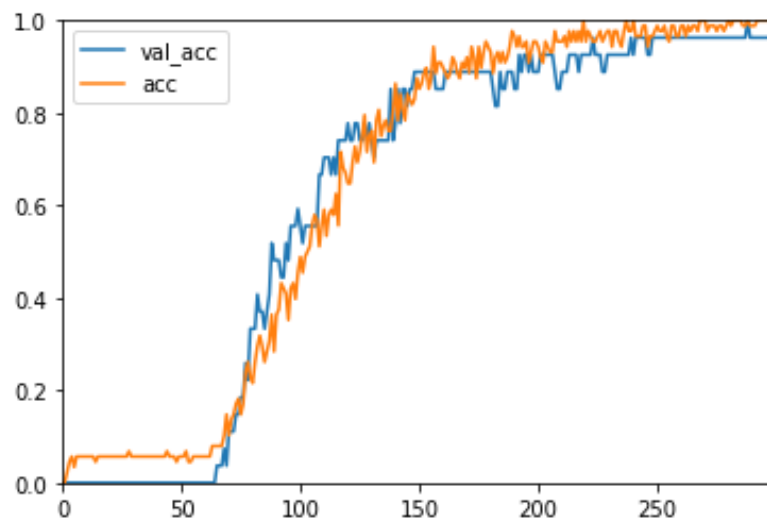


Fig. 8. Graphical delegation of Average Accuracy of model using GRU and Word2Vec with 400 Epochs.

Figure:- 8 illustrates the graphical representation of our model which gives the accuracy of 98.86% with the combination of GRU and word2Vec word embedding.

Chapter 5

Conclusion and Future Work

In this work, we effectively developed and trained heuristic neural framework applied on Bangla language that learns from question-answer pair with supporting fact and applicable to answer several paraphrased questions having single answer. This model with several contributions slightly approaches better execution using word2vec sentence representation. We have also experimented positioning encoding as sentence and question representation in the time of training the model which performs less accuracy than pre-trained Bangla word2vec with 300 dimensions vector space representing each word of the Bangla corpus. By changing optimizer of GRU and LSTM based RNN variation on consequence have been come out.

However, as working with Bangla dataset is quit exhausting and originating Bangla question manually is tedious, our work could get more further improvement by applying approach on a large question-answer dataset considering: chaining fact, induction, deduction etc. tasks. So, still we have much to work so for the further enhancement of complete QA system. Furthermore, our model is still unable to answer without supervision fact which is also a field of working in future.

Bibliography

- [1] Iyyer, M., Boyd-Graber, J., Claudino, L., Socher, R., & Daumé III, H. (2014, October). A neural network for factoid question answering over paragraphs. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 633-644).
- [2] Dong, L., Mallinson, J., Reddy, S., & Lapata, M. (2017). Learning to paraphrase for question answering. *arXiv preprint arXiv:1708.06022*
- [3] Weißenborn, Dirk & Wiese, Georg & Seiffe, Laura. (2017). FastQA: A Simple and Efficient Neural Architecture for Question Answering.
- [4] Iyyer, M., Yih, W. T., & Chang, M. W. (2017, July). Search-based neural structured learning for sequential question answering. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1821-1831).
- [5] Weston, J., Bordes, A., Chopra, S., Rush, A. M., van Merriënboer, B., Joulin, A., & Mikolov, T. (2015). Towards ai-complete question answering: A set of prerequisite toy tasks. *arXiv preprint arXiv:1502.05698*.
- [6] Sukhbaatar, S., Weston, J., & Fergus, R. (2015). End-to-end memory networks. In *Advances in neural information processing systems* (pp. 2440-2448).
- [7] J. Chung, C., . Gülc ehre, K. Cho, and Y. Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint: 1412.3555*, 2014.
- [8] A. Graves. Generating sequences with recurrent neural networks. *arXiv preprint: 1308.0850*, 2013.

- [9] T. Mikolov, A. Joulin, S. Chopra, M. Mathieu, and M. Ranzato. Learning longer memory in recurrent neural networks. arXiv preprint: 1412.7753, 2014
- [10] A. H. Mirza, "Variants of combinations of additive and multiplicative updates for GRU neural networks," *2018 26th Signal Processing and Communications Applications Conference (SIU)*, Izmir, 2018, pp.1-4. doi: 10.1109/SIU.2018.8404457
- [11] E. Nurvitadhi, Jaewoong Sim, D. Sheffield, A. Mishra, S. Krishnan and D. Marr, "Accelerating recurrent neural networks in analytics servers: Comparison of FPGA, CPU, GPU, and ASIC," *2016 26th International Conference on Field Programmable Logic and Applications (FPL)*, Lausanne, 2016, pp.1-4. doi: 10.1109/FPL.2016.7577314
- [12] Alam, Firoj & Chowdhury, Shammur & Noori, Sheak. (2016). Bidirectional LSTMs — CRFs networks for bangla POS tagging. 10.1109/ICCITECHN.2016.7860227.
- [13] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems* (pp. 3111-3119).

M.Uddin,N.S.Patwary, M.Hasan, T.Rahman, “End-To-End Neural Network for Paraphrased Question Answering Architecture with single supporting line in Bangla Language” in *Proceedings of the 2020 International Conference on Signal Processing and Information Communications (ICSPIC)*. Bali, Indonesia, Jan. 2020