e-ISSN : 0976-5166
p-ISSN : 2231-3850

AUTOMATIC QUESTION ANSWERING IN MYANMAR LANGUAGE USING SEQUENCE TO SEQUENCE WITH ATTENTION MECHANISM

[1]Naing Naing Khin

Natural Language Processing Lab, Faculty of Computer Science
University of Computer Studies, Yangon
Myanmar
Email:naingkhin@ucsy.edu.mm

[2]Khin Mar Soe

Natural Language Processing Lab, Faculty of Computer Science
University of Computer Studies, Yangon
Myanmar
Email:khinmarsoe@ucsy.edu.mm

Abstract

The primary and enduring objective of natural language processing and artificial intelligence is teaching machines to read and comprehend human natural language. Question Answering is a conversational model which can generate an answer automatically from a given context. In many systems, they have great potential for those interested in learning the necessary details. QA can be based on various techniques, including information retrieval, knowledge-based, generative, and rule-based approaches. Sequence to Sequence (Seq2Seq) model is adapted for good conversational model for QA. There are many Question Answering Models in English but there is still need a system that can provide convenient question and answer in Myanmar Language. In this research, we proposed Automatic Question Answering in Myanmar language utilizing the RNN encoder-decoder model-based Sequence to Sequence model with Attention Mechanism. 34,850 Myanmar language question-answer pairs are used to train the question answering model. The experimental result is evaluated on Precision, Recall, F1 Score, Human Evaluation and gets overall accuracy 71%.

***Keywords*: Recurrent Neural Network(RNN); Sequence to Sequence learning; Attention Mechanism; Natural Language Processing(NLP), Question Answering.**

1. Introduction

In Natural Language Processing(NLP), Question Answering with deep learning is a crucial and challenging task. Deep Learning methods, there has been a revolution in natural language processing in general as well as machine understanding. They have significantly outperformed all conventional NLP and machine learning techniques, and they are actively propelling the field's expansion. Question Answering systems are greatly striven after in many areas of industry. The two domain types on which the Question Answering System operates are Close Domain System and Open Domain System. Closed-domain question answering is mostly utilized for task-oriented systems and addresses inquiries specific to a given topic. Open Domain Systems are not limited to any one domain, and a vast amount of data is required in order to provide the user with the appropriate response. The benefits of computers being able to understand language and reason could be huge. Various types of real-life applications of QA commercial systems such as Google Now, Siri, and Alexa are being utilized extensively as speech recognition technology advances to human levels. Furthermore, one of the main reasons text understanding by machines is such a crucial issue is because of all the text material that is readily available online. With the growth of Question Answering systems are key to language development, they are widely used ranging from restaurant recommendations to medical diagnosis. Most research in Natural Language Processing is conducted in English Language, it is difficult to say how well the current state-of-the art models would perform when presented with Myanmar text data because of data scarcity in low-resourced language. In this study, we have proposed a Myanmar Language Question Answering using Neural Network approach with an attention mechanism based on the Sequence to Sequence model. Neural networks can handle complex tasks more accurately, such as facial recognition, document summarization, text generation and so on.

2. Related Works

In a variety of contexts and applications, such as search engines, virtual assistants, chatbots for customer service, and educational platforms, Question Answering Systems seek to mimic human cognitive abilities to understand queries and offer insightful responses. The design of the QA system and the kinds of issues it is intended to address will determine the particular techniques employed in each stage and the architecture of the system.

Language model embeddings with a BERT pre-trained language model is used the Knowledge Base Question Answering task in Sharath et al., [18]. It is based on Convolutional Neural Network architecture with Multi-Head Attention mechanism to represent the asked question in multiple aspects. Experiments were conducted on the Freebase Knowledge Base and the Web-Questions dataset. The experimental result reaches for the average F1 score is 52.7%.

The Question Answering system Bansal et al., [4] is created based on encoder decoder structure that can extract information about java methods directly from raw source code. In this research, experiment is conducted with human users to evaluate the system depend on relevance, accuracy, completeness and conciseness. It describes 79% of model responses are correct. By comparing the model's accuracy with other machine learning model evaluation techniques, it could be more accurate result.

Xiang et al., [22] suggested a system for answering questions based on the self-attention mechanism and bi-LSTM. The self-attention process enhances the sentence's textual characteristics and semantic content. When processing lengthy information sequences, the self-attention mechanism can ignore the length limitation in contrast to the traditional attention mechanism. It proves the viability and efficacy of the method by obtaining a final accuracy rate of 65.5% from the model answer.

Attention-based Recurrent Neural Networks for Question Answering by Hong et al., [13] explored and compared two of the models Match-LSTM and Bidirectional Attention Flow (BiDAF) and proposed an ensemble model which combined these two models. The comparison visualization demonstrates that every model has its own strengths and weaknesses, while the outcomes visualization demonstrates that the model is correctly applying attention and calculating. The highest F1 score for BiDAF model achieves 63.63% and EM score is 40.80% respectively.

Day et al., [9] delved a factoid question-answering system architecture that combines the answer type categorization (AT) and question expected answer type (Q-EAT) models with BERT. In order to investigate whether classification will have an impact on question answering systems' prediction outcomes, the two classification models are integrated with question answering models in this study. The question answering and classification systems were optimized using the BERT pre-training model. The proposed model results the F1 score 89.44% and Exact Match 80.19%. The prediction accuracy EM of the question answering system will be enhanced when the query sentence and the answer categorization match.

3. Question Answering in AI and NLP

Creating effective Question-Answering (QA) systems is a major undertaking in the rapidly increasing natural language processing field. These kinds of systems have a ton of potential uses, from chatbots for customer support to research assistance. In order to accomplish the human interaction with natural language, engineers and researchers in AI have created a unique kind of model called the Question-Answering model. This model takes in a question and then process a large amount of text data to determine the most accurate answer. A question answering model's ultimate objective is to accurately comprehend the question's meaning and offer a response that makes sense in the given situation. A QA model's effectiveness is determined by how well it can respond to a variety of queries with precise and insightful information. The need for virtual assistants and conversational AI systems has increased recently, which has fueled the field of natural language processing's advancement in question answering. These systems analyze queries, locate pertinent information, and produce responses using advanced natural language processing techniques. These days, answering questions in natural language with OpenAI is beneficial and popular. The process of developing an NLP question-answering system is made simpler, cheaper, and less complex. If we have sufficient data sources, anyone can develop an efficient natural language-based quality assurance system with the correct resources and process knowledge. NLP question-answering systems with OpenAI integration can boost output and enhance client satisfaction.

4. Challenges of Question Answering in Myanmar Language

Answering questions is a challenging task because it requires overcoming several obstacles in order to function properly. Despite the question's complex or ambiguous wording, the first issue is to discern its intended meaning. This necessitates a profound comprehension of real language and the capacity for the AI system to distinguish between words and phrases that sound same. Especially in Myanmar(Burmese) Language, due to the abundance of English borrowing words, answering questions in this language poses a special difficulty. The first challenge is to understand the intention behind the question, even if the wording is complicated or unclear. The question in English “Where is your university located?”: “မင်းရဲ့တက္ကသိုလ် က ဘယ်မှာ တည်ရှိ တာလဲ” and “What is the address of your University?”: “မင်း တက္ကသိုလ် ရဲ့လိပ်စာ က ဘာလဲ” meanings are the same semantic in Myanmar and the answer for these two questions can be the same. But the similarity for comparing these two sentences are different. The second challenge is lack of question-answer data source in Myanmar Language because of low resource language. This study is the first Myanmar Question Answering in Neural Network based approach and still need some demands in text processing.

5. Generating Text with Sequence to Sequence Model

One of the most popular neural network models for sequential learning nowadays is the sequence to sequence model, which is used in speech recognition, image captioning, dialogue systems, and other systems. It is made up of an encoder and a decoder, two Recurrent Neural Networks (RNN): one each for the decoder and encoder. The source sentence is read by the encoder RNN, and the final state is utilized as the decoder RNN's initial state. Every piece of information about the source is encoded in the final encoder state, and the decoder can use this vector to produce the target sentence. The input sequence for a question-answering problem consists of all the words in the query. The input sequence is fed into the encoder, which transforms it into a fixed-size feature vector [17]. The neural seq2seq model needs to keep track the word that is currently being processed by its encoder or decoder in order to generate text. It accomplishes this by using timesteps, each of which shows the token in a given sentence that the model is presently handling.

5.1. *Attention Mechanism*

To improve the performance of the RNN model, which encodes the input sequence into a single fixed size vector, an attention mechanism has been designed. Large sequence decoding is a challenge for the simple seq2seq. The amount of information in a sequence may decrease as it gets longer. In this study, we incorporate the attention mechanism, whereby the model attempts to predict an output word by utilizing only those portions of an input where the most pertinent information is concentrated, rather than utilizing the complete sentence. Instead of applying a fixed context, a different context vector is used for generating word [7], [1], [14]. As an encoder and decoder, two-layered RNNs with an LSTM cell are employed. Because it includes both forward and backward directions, this model can calculate both previous and next words. In order to predict or infer a single element, such as a pixel in an image or a word in a sentence, attention in deep learning can therefore be understood as a vector of significant weights. The target element is approximated by taking the sum of its values weighted by the attention vector. We do this by estimating the attention vector's correlation strength with other elements. The attention mechanism can assist in increasing prediction accuracy and streamlining the model's processing speed by enabling it to concentrate on the most pertinent facts. Due to their capacity to dynamically focus on various input segments, seq2seq models with attention typically perform better than regular seq2seq models on tasks involving lengthy and complex sequences. The attention mechanism's alignment makes it simpler to analyze and comprehend the actions and judgments of the model. This is especially helpful for jobs like translation, where it's critical to determine which input words translate into which output words.

5.1.1. Word Embedding with Attention

Words can be represented densely in vector form in a continuous vector space using word embeddings. FastText, GloVe, and Word2Vec are popular methods for creating word embeddings. Through the conversion of textual input into a format that machine learning models can handle, these embeddings serve as the basis for numerous NLP activities. To build an effective sequence-to-sequence model, Word Embedding Layer is essential in converting input words into dense vectors using an embedding layer. Next, the model processes the embedded input sequence through an Encoder (e.g, LSTM or GRU) to obtain hidden states. The attention mechanism is then applied to compute context vectors for the decoder. Finally, the Decoder uses the context vectors and the previous decoder hidden states to generate the output sequence.

5.1.2 Attention-based Model

There are many attention-based models, among them, two types of categories are global and local attention. Both approaches first take as input the hidden state h_t at the top of a stacking RNN. Next, in order to assist in predicting the current target word y_t, the objective is to derive a context vector, c_t, that gathers pertinent source-side data. These models have different methods for deriving the context vector (c_t), but they all follow the same procedures after that. To calculate the alignment scores, an alignment model makes use of the annotations and the decoder's current hidden state:

$$a_t(s) = align(h_t, h_s^-) = \frac{\exp(score(h_t, h_s^-))}{\Sigma_{s'} \exp(score(h_t, h_{s'}^-))} \quad (1)$$

, where h_t is the target hidden state and h^-_s is the source hidden state.

Three different score functions are calculated to compute attention scores: dot, general and concat methods. Alignment scoring functions calculate similarity between the source and target hidden states.

$$score(h_t, h_s^-) = h_t^T h_s^- \quad \text{(dot)} \quad (2)$$

$$score(h_t, h_s^-) = h_t^T W_a h_s^- \quad \text{(general)} \quad (3)$$

$$score(h_t, h_s^-) = \upsilon_a^T \tanh(W_a[h_t; h_s^-]) \quad \text{(concat)} \quad (4)$$

While the concept put forth by Bahdanau et al., [3] and the global attention approach are conceptually similar, there are a few significant variations that show how we have both simplified and generalized from the original model. First, as shown in Figure 1, we just employ hidden states at the top RNN layers in the encoder and decoder. When determining the context vector (c_t), a global attentional model considers all of the encoder's hidden states. The current target hidden state, h_t, is compared with each source hidden state, h^-_s, to create a variable-length alignment vector a_t, whose size equals the number of time steps on the source side in this model type. The weights are used to generate a context vector c_t through the weighted sum as:

$$c_t = \Sigma_{i=1}^T a_{t,i} h_s^- \quad (5)$$

An attentional hidden state is computed based on a weighted concatenation of the context vector and the current decoder hidden state:

$$h_t^\sim = \tanh(W_c[c_t; h_t]) \quad (6)$$

After passing through the softmax layer with the attentional vector $h_t^\sim$, the decoder produces the final output :

$$y_t = Soft\max(W_y h_t^\sim) \quad (7)$$

In figure 1, the source sentence "မေရီ အင်္ဂလိပ်စကား ပြောတတ် သလား" "Can Mary speak English?" is entered as an input sentence and the input word embedding is being processed. The alignment score is calculated between source and target hidden states. And, apply the softmax function on the alignment score effective transforming them into weight values. The weights are applied in order to create a context vector by the weighted sum. And the attention hidden state pass through the decoder and produces the final output: "ပြောတတ် တယ်" "Can speak" as result.

The architecture of attention mechanism for the question answering model is described as follow:

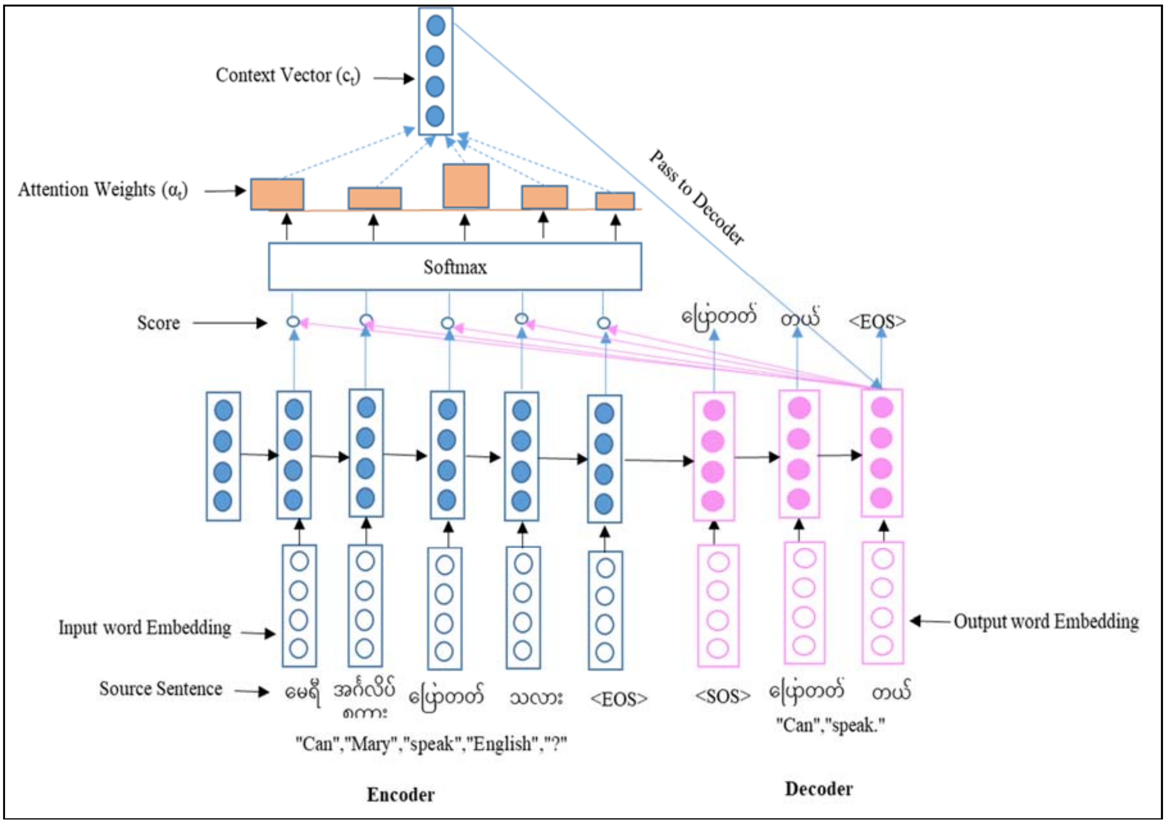

Fig. 1. Attention in Question Answering Model

6. Proposed System Architecture

There are two stages to the proposed system: training and testing. Prior to training the question answering model, the data is collected as Question-Answer pairs from the available data sources and makes preprocessing. After preprocessing, the model is then trained using both sequence-to-sequence (based line) and attention-based sequence-to-sequence. Difference score functions make the model accuracy changes: dot, general and concat methods are used to train the question answering model. F1 score, accuracy, precision, recall, and, and human evaluation techniques are used to generate the response and assess the model. In testing phase, the input for a user question is preprocessed before feeding into the model and test the system to generate the answer using question answering model.

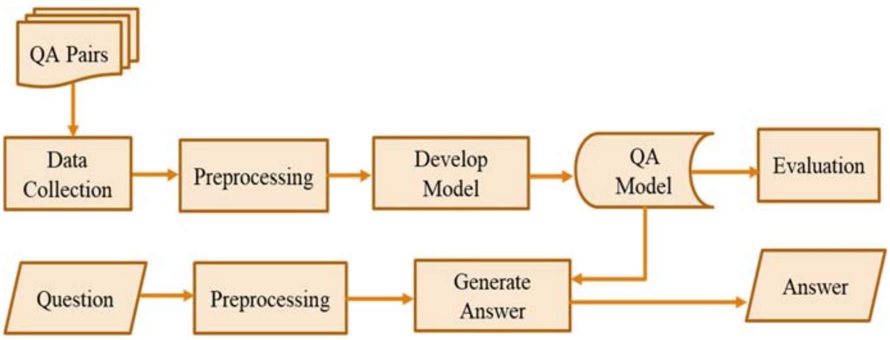

Fig. 2. Process Flow of Proposed Question Answering Model

6.1. *Data Collection*

When developing a system for natural language processing applications, data is the most crucial component. However, it is also a laborious effort, particularly for languages with limited resources. Language-specific processing and a substantial amount of training data are required for most NLP models. The language of Myanmar is likewise under resourced, and we have trouble finding enough data to create training sets. To train the question answering model, we have gathered a corpus of over 30k question-answer pairs. The data that has been cited from the Stanford Question Answering Dataset (SQuAD). SQuAD is used to test and train question-answering algorithms. It comprises of actual questions that people have asked on a collection of Wikipedia pages, with each question's response being a particular passage from the related articles.

6.2. *Data Preprocessing*

The data from the SQuAD dataset is translated into Burmese and attached to a set of questions and answers. Then check manually because English and Burmese language are different in nature. Segmentation can be particularly difficult in languages without clear boundaries such Japanese, Chinese and also Myanmar Language. So, segmentation is needed to separate the meaningful word and phrases. Segmenting the collected data into word is processed by using Myanmar word segmenter tool of UCSY NLP lab. Myanmar Word Segmentation employs the longest matching and bigram method of segmentation along with a pre-segmented corpus. The sentence "ဘီယွန်းစေးကဘယ်ဆယ်စုနှစ်မှာနာမည်ကြီးလာတာလဲ" "When did Beyonce become popular?" in English to "ဘီယွန်းစေး က ဘယ် ဆယ်စုနှစ် မှာ နာမည် ကြီး လာ တာလဲ". Pre-processing of source and query texts, like punctuation symbols removal, lowercasing the English letter and manually corrected the numbers. Remove extra whitespaces, including leading, trailing, and multiple spaces between words. Eliminate non-alphanumeric characters such as Emojis that are not useful for the task. These preparation procedures guarantee that the Unicode dataset is ready for a range of natural language processing applications, which produces more accurate and dependable models.

7. Experimental Result

Our research has contributed to the creation of a corpus of Myanmar language and the improvement of automatic question answering using the sequence-to-sequence model. Through the attention mechanism, the system employs an encoder decoder model based on recurrent neural networks. In order to address the issue of lengthy sequences which arises in the majority of NLP tasks: attention was brought to the table. We have discovered that the baseline RNN encoder decoder approach is not as good as the result obtained with the attention mechanism. The model's overall accuracy is not particularly great, and both the model and the data preparation still require some work. The Pytorch library and the Python programming language were utilized to process the system's implementation.

Question:	ဗုဒ္ဓဘာသာ က ဘယ်လို ဘာသာ မျိုးလဲ	What kind of religion is Buddhism?
Answer:	အကြောင်း အကျိုး တရား ပေါ်မှာ အခြေခံ သော ဘာသာ	A religion based on cause and effect
Question:	ကမ္ဘာမြေ ၏ သမုဒ္ဒရာ များ သည် ကမ္ဘာ ၏ မည်သည့် ရာခိုင်နှုန်း ကို လွှမ်းခြုံ ထား သနည်း	What percentage of the Earth's oceans cover?
Answer:	၇၁ ရာခိုင်နှုန်း	71 percentage
Question:	အပူ ဒြပ်ထု ဆိုတာ ဘာလဲ	What is thermal mass?
Answer:	အပူ သိုလှောင်နိုင် စုပ်ယူနိုင် ထုတ်လွှတ်နိုင်ခြင်း ကို ခေါ်ဆိုသည်	It is called the ability to store and absorb and release heat
Question:	အန္တာတိကတိုက် မှာ ဘယ်လို ပင်ဂွင်း မျိုးစိတ် တွေ နေထိုင် လဲ	What species of penguins live in Antarctica?
Answer:	ဧကရာဇ် ပင်ဂွင်း	The emperor penguin
Question:	ဘယ် စနစ် က နှလုံးခုန် နှုန်း ကို ထိန်းချုပ် ပေး တာလဲ	Which system controls heart rate?
Answer:	အလိုအလျောက် အာရုံကြောစနစ် နှစ်ခု	Two autonomic nervous systems

Table 1. Sample Output of Question Answering

7.1. *Evaluation of Question Answering Model*

Tested the system by the two models of seq2seq model and seq2seq with attention mechanism. While both these models aim to map input sequences to output sequences, the addition of the attention mechanism significantly enhances the model's ability to handle longer sequences, capture finer details, and provide better performance and interpretability. Random 2000 questions are used in this testing as user input questions and the resulted answer is measured by Accuracy, Precision, Recall, F1 score. Precision in a QA system indicates how many of the answers provided are correct, focusing on the accuracy of the results. Recall measures how many of the correct answers the system successfully identifies, focusing on completeness. F1 score offers a balanced measure, considering both precision and recall. The models were trained using learning rate 0.0001, hidden size 512, batches of 64, 128 and 256, respectively.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (8)$$

$$\Pr ecision = \frac{TP}{TP + FP} \quad (9)$$

$$\operatorname{Re} call = \frac{TP}{TP + FN} \quad (10)$$

$$F1 = \frac{2 * \Pr ecision * \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call} \quad (11)$$

		Accuracy	Precision	Recall	F1
Seq2Seq(Based Line Model)		0.60	0.64	0.59	0.61
Seq2Seq with Attention	Dot	0.71	0.72	0.74	0.73
	Genenal	0.61	0.62	0.64	0.62
	Concat	0.63	0.67	0.69	0.67

Table 2. Evaluation of Precision, Recall, F1 Score

Table 2 describes the evaluation score for the sequence to sequence model (based line) and sequence to sequence with attention mechanism with three different score functions: dot, general and concat methods. Among them Seq2Seq with Attention(Dot) method gets the highest accuracy.

The accuracy of the trained models is shown in figure 3. An attention-based model outperforms based line model in terms of score. These two models' score comparisons are illustrated as follows:

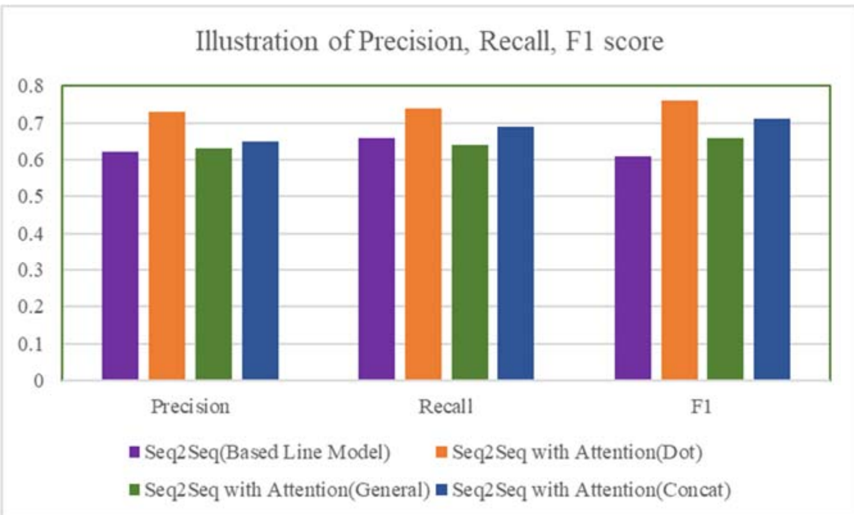

Fig. 3. Changes in Evaluation of Proposed Question Answering Model

7.2. *Human Evaluation*

Evaluating the quality and relevance of QA systems' answers can also be done by using evaluation methods that involve human judges, experts, or users. Manual evaluation requires human evaluators to rate the system's answer based on criteria such as accuracy, completeness, clarity, and relevance. As an alternative, user behavior and satisfaction approaches examine how satisfied users are with the responses provided by the system. These methods can offer more contextual and qualitative insights about the functionality and performance of the system.

Question	Referenced Answer	Predicted Answer
၂၀၁၀ မှာ ဘာစီလိုနာ ရဲ့ဝင်ငွေ က ဘယ်လောက် လဲ	ယူရို ၃၆၆ သန်း	ယူရို ၃၆၆
အင်တာနက် အာခိုင့် ရုံး ချုပ် က ဘယ် မှာလဲ	ဆန်ဖရန်စစ္စကို	ဆန်ဖရန်စစ္စကို မှာ ဆန္ဒပြ
ဘီယွန်းစေး ရဲ့ပထမဆုံး တစ်ကိုယ်တော် အယ်လ်ဘမ် နာမည် က ဘာလဲ	အချစ် ၌ အန္တရာယ် ရှိသည်	အချစ် ၌ အန္တရာယ်
ပိုလီ နည်းပညာ ကျောင်း တွေ ကို တက္ကသိုလ် တွေ ဖြစ် လာ အောင် ဘယ် လုပ်ရပ် က ခွင့် ပြု ခဲ့လဲ	၁၉၉၂ ခုနှစ် ထပ်ဆင့် နှင့် အဆင့်မြင့် ပညာရေး ဥပဒေ	ဘာစီလိုနာ အသင်းသား များ
ဗုဒ္ဓဘာသာ ဝင်များ လိုက် နာ ရ မည့် အခြားသော ကျင့်ထုံးများ က ဘာလဲ	ကုသိုလ် ဆယ် ပါး	ကုသိုလ် ဆယ် ပါး

Table 3. Human Evaluation for the proposed model

In table 3, it has showed the result of model predicted answer and the original answer. As human evaluation, some outputs are not completed answer but it can assume acceptable with semantic in language. For example, "အချစ် ၌ အန္တရာယ်" in the predicted answer is correct although the original referenced answer is "အချစ် ၌ အန္တရာယ် ရှိသည်". So, Human evaluation is necessary for some situations.

8. Conclusion

As information in day to day life is increasing, Question Answering system is an indispensable part in the communication process. QA system allows a user to express in natural language and get an immediate and brief response. In this research we have designed an automated Question Answering for Myanmar Language utilizing an attention-based sequence-to-sequence methodology. Comparing the output of the QA system with a reference answer or a set of approved answers can help assess the relevance and quality of the system's responses. This is the first attempt of Myanmar Language Automatic Question Answering System and have to develop with more data and techniques. We have trained the QA model on Tesla k80 GPU, NVIDIA GeForce RTX 4060 and implemented by Python language with Pytorch library. The experimental results showed some insights for the question-answering system are provided by the Sequence-to-Sequence model with attention mechanism. There are still some challenges in question answering such as domain-specific understanding, handling ambiguous queries, and ensuring ethical deployment of these systems remain areas of active research and development. We have continued our work with Transformer Model for next step of our research.

Acknowledgments

Throughout my study, my committed supervisor has provided constant direction, inspiration, and steadfast support. Her encouragement, and expertise were instrumental in bringing this research to fruition. I am deeply grateful for their generosity and unwavering belief in this endeavor provided invaluable insights, constructive feedback, and encouragement at every stage of my research journey.

References

[1] Ali, A.; Amin, M. Z. : Conversational AI Chatbot Based on Encoder Decoder Architectures with Attention Mechanism, Nov 2019.
[2] Arora, R.; Singh, P.; Goyal, H.; Singhal, S.; Vijayvargiya, S. : Comparative Question Answering System based on Natural Language Processing and Machine Learning, Proceedings of the International Conference on Artificial Intelligence and Smart Systems (ICAIS-2021).
[3] Bahdanau, D.; Cho, K.; Bengio, Y. : Neural Machine Translation By Jointly Learning To Align And Translate, 2015.
[4] Bansal, A.; Eberhart, Z.; Wu, L.; McMillan, C. : A Neural Question Answering System for Basic Questions about Subroutines, 11 Jan 2021.
[5] Bhuiyan,R.; Masum, A.; Oaphy, A.; Hossain, S.; Abujar, S. : An Approach for Bengali Automatic Question Answering System using Attention Mechanism, 11th ICCCNT 2020.

[6] Bi, M.; Zhang, Q.; Zuo, M.; Xu, Z.; Jin, Q. : Bi directional LSTM Model with Symptoms Frequency Position Attention for Question Answering System in Medical Domain, 2019.
[7] Chandra, Y. W.; Suyanto, S. : Indonesian Chatbot of University Admission Using a Question Answering System Based on Sequence-to-Sequence Model , 2019.
[8] Costa-jussa, M. R.; Nuez, A; Segura, C. : Experimental Research on Encoder-Decoder Architectures with Attention for Chatbots, Vol. 22, No. 4, 2018.
[9] Day, M; Kuo, Y : A Study of Deep Learning for Factoid Question Answering System, 2020.
[10] Dhyani, M.; Kumar, R. : An intelligent Chatbot using deep learning with Bidirectional RNN and attention model, Volume 34, Part 3, 2021, Pages 817-824.
[11] Galassi, A.; Lippi, M.; Torroni, P. : Attention in Natural Language Processing, May 2020.
[12] Gennaro, G. D.; Buonanno, A.; Girolamo, A. D.; Ospedale, A.; Palmieri, F. A. N. : Intent Classification in Question-Answering Using LSTM Architectures, 25 Jan, 2020.
[13] Hong, D.; Wan, B. : Attention-based Recurrent Neural Networks for Question Answering, 2016.
[14] Luong, M.; Pham, H.; Manning, C. : Effective Approaches to Attention-based Neural Machine Translation, Sep 2015.
[15] Mutabazi, E.; Ni, J.; Tang, G.; Cao, W. : A Review on Medical Textual Question Answering Systems Based on Deep Learning Approaches, Appl. Sci. 2021.
[16] Nie, Y.; Han, Y.; Huang, J.; Jiao, B.; Li, A. : Attention-based encoder decoder model for answer selection in question answering, March 2017.
[17] Palasundram, K.; Sharef, N. M.; Kasmiran, K.A.; Azman, A. A. : Enhancements to the Sequence-to-Sequence-Based Natural Answer Generation Models, March 2020.
[18] Sharath, S.; Rekabdar, B.: Question Answering over Knowledge Base using Language Model Embeddings, 17 October 2020.
[19] Sutskever, I.; Vinyals, O. ; Le, Q. V. : Sequence to Sequence Learning with Neural Networks, 2014.
[20] Van-Tu, N.; Anh-Cuong, L. : Improving Question Classification by Feature Extraction and Selection, May 2016.
[21] Wang, W. : Self-Attention in Question Answering, 2016.
[22] Xiang, H; Gu, J : Research on Question Answering System Based on Bi-LSTM and Self-attention Mechanism, 2020.
[23] Zhang, B.; Wang, H.; Jiang, L.; Yuan, S.; Li, M. : A Novel Bidirectional LSTM and Attention Mechanism Based Neural Network for Answer Selection in Community Question Answering, CMC, vol.62, no.3, pp.1273-1288, 2020.

Authors Profile

In 2010, Naing Naing Khin earned her Master of Computer Science (M.C.Sc.) degree from the University of Computer Studies, Sittway, Myanmar. She is currently pursuing her Ph.D. at the University of Computer Studies, Yangon, Myanmar, and is a member of the Natural Language Processing and Speech Processing Lab at UCSY. Her research interests are Natural Language Processing, Machine Learning, and Image Processing.

In 2005, Dr. Khin Mar Soe finished her doctoral studies in information technology. She is a professor at the University of Computer Studies, Yangon, where she also leads the Natural Language Processing Lab at the Faculty of Computer Science. In the area of natural language processing, she actively supervises Ph.D. and Master's theses. Dr. Khin Mar Soe has also participated to the Asia Language Treebank (ALT) project, which includes research collaboration between NICT, Japan, and UCSY, and ASEAN MT, a machine translation program for South East Asian languages.