Generating Myanmar News Headlines using Recursive Neural Network

Yamin Thu
Natural Language Processing Lab
University of Computer Studies, Yangon
Yangon, Myanmar
yaminthu@ucsy.edu.mm

Win Pa Pa
Natural language Processing Lab
University of Computer Studies,Yangon
Yangon, Myanmar
winpapa@ucsy.edu.mm

Abstract

Text summarization in the form of Headline prediction for written articles becomes a popular research recently. This paper presents a headline prediction model using Recursive Recurrent Neural Network (Recursive RNN) for Myanmar articles and evaluates the performance by comparing with sequence-to-sequence models. Recursive RNN model use encoder-decoder LSTM model that generates a single word forecast and call it recursively to text summarization in the form of news headlines for Myanmar news articles. A Recursive RNN, which takes input as the sequence of variable length and has to generate the variable length output by taking into account the previous input. 5000 articles of Myanmar news were collected to train headline prediction model. The performance of Recursive RNN, Seq2Seq with one-hot encoding and Seq2Seq with word embedding (GloVe) were evaluated in terms of ROUGE score values. The experimental results show that Recursive RNN model significantly better than other two models.

Keywords—*LSTM, Recursive RNN, ROUGE, word embedding*

I. INTRODUCTION

Huge amount of information are available on the internet news websites but users want to extract best particular information within a short search time. Document summarization has two types extractive and abstractive. Extractive method chooses the salience sentences from the original document. Abstractive method uses the encoder-decoder architecture which generates latent factor representation of the data and decodes it to generate a summary1. Abstractive model can be more effective by performing generation from scratch but slow and inaccurate encoding of very long documents [9]. Recursive RNN-based text summarization model was popular generating a new headline or short summary consisting of a few sentences that extract the most important point of article or a passage. Recursive RNN based summarization model to be very effective for transforming Unicode Myanmar text from one to another. These models are trained on large amount of input sequence of Myanmar words and expected output sequence of Myanmar words and generate output sequences based on the given input. Text summarization is the most challenging problem in Myanmar Language. There are many automatic text summarizers for other languages but a few for Myanmar language. Therefore, as our motivation, we was collected Myanmar news corpus including title of Myanmar news articles and description of text using Myanmar standard Unicode font in Myanmar news web pages. Three contributions are described in this paper. The first one is obtain the context vectors from the Myanmar news document which can be used for extractive summarization task. The second one is using RNN headline training model for Myanmar news summarization. The third one is using document context vector, produce the Myanmar news summarization. The rest of the paper is organized as follows: section 2 describes related work and section3 describes challenges of Myanmar news summarization section4 describe the overview of RNN-based Myanmar text summarization. Section 5 experimental setups and evaluation results and section 6 discusses conclusion and future work direction.

[1] https://mc.ai/extractive-and-abstractive-text-summarization-methods/

II. RELATED WORK

The author proposed abstractive text summarization using Attention Encoder-Decoder Recurrent Neural Networks and modeling key-words. In this paper contain capturing the hierarchy of sentence-to-word structure, and producing rare word or unseen word at training time [1]. Khatri, Chandra, Gyanit Singh, and Nish Parikh proposed Document-context Sequence to sequence (Seq2Seq) learning has recently been used for abstractive and extractive summarization. It gives humans high-level understanding of the document and Seq2Seq model will generate much richer document specific summaries [2]. In this paper, using headline generation models, train the English GigWord corpus (Graff & Cieri, 2003), consisting of news articles from number of publishers. Summarize the first sentence of an article is used to predict the headline, using RNN model, RNN-based encoder-decoder models with attention (seq2seq) perform well ROUGE (Lin, 2004), an automatic metric often used in summarization, and human evaluation [3]. Present Recursive Neural Networks (R2N2) for the sentence ranking task of multi-document summarization. It transforms the ranking task into a hierarchical reversion process which is modeled by recursive neural networks [4]. In this paper describe about the performance of LSTM bidirectional and Seq2Seq model on Amazon reviews and news dataset is compared using BLEU, ROUGE_1 and ROUGE_2 score [5]. In the paper [6] describe about a fully data-driven approach to abstractive sentence summarization. This technique useful for local attention-based model that generates each word of the summary based on the source sentence. Although the model is simple, it can be easily trained on a large amount of training dataset. This paper emphasized on feed-forward neural network with attention-based encoder to resolve the challenge of abstractive summarization. Briefly explain attentive recurrent neural network and recurrent neural network encoder-decoder. Amazon reviews dataset and Bidirectional LSTM model using Sequence to Sequence model compared with the human generated summary [7].

III. NATURE AND COLLATION OF MYANMAR LANGUAGE AND SEGMENTATION

The nature of Myanmar script has 33 consonants and the consonants combine with vowels and medial to form the complete syllable in Myanmar language. Myanmar language does not place spaces between words but spaces are usually used to separate phrases. In Myanmar language, sentences are clearly delimited by a sentence boundary marker pote ma (။) but words are not always delimited by spaces. In Myanmar language segmentation process is the fundamental and very important step in generating the summaries. Segmentation process includes sentence and word segmentation that are finding boundaries of sentences and word tokens. In English, word boundaries can be easily defined but it is not easy in Myanmar. Myanmar words are written continuously without using space in sentences. Therefore word segmentation process is very important phase to pass word tokens and sentences to next steps of summarization. For Myanmar sentence and word segmentation, we used Myanmar word segmenter of UCSY [8]. For example,

Input Myanmar Sentence:
ပြင်ဦးလွင်မြို့ဒီဇင်ဘာခြံတွင်ပန်းအလှဆင်ပွဲတော်ကျင်းပမည်။
Segmented Sentence:
ပြင်ဦးလွင်မြို့_ ဒီဇင်ဘာ_ ခြံ_ တွင်_ ပန်း_ အလှ_ ဆင်_ ပွဲတော်_ ကျင်းပမည်_ ။

The overview of the proposed system described in figure 1. The proposed system used the concept of machine learning based Recursive RNN predicted headline summary. And then preprocessing steps are performed. In preprocessing step, the system performs word segmentation, stop words removal such as pronouns, prepositions, conjunction and particles etc.

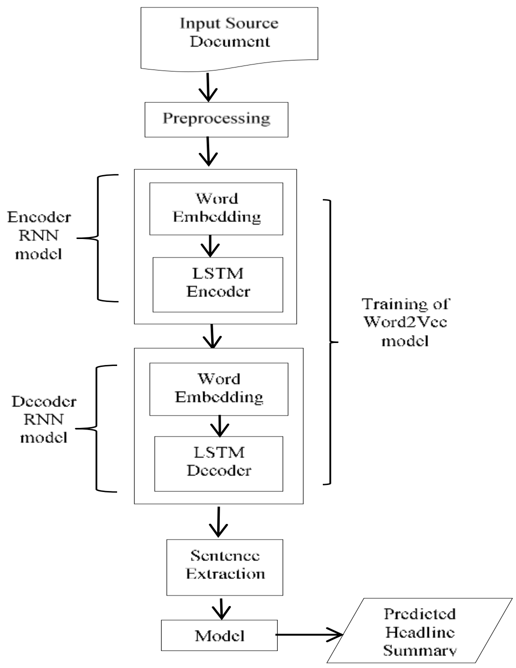

Figure 1. Recursive RNN-based Summarization model overview

And then Headline Generator is using the Recursive RNN-based model and consists of encoder and decoder model, the encoder use one RNN and the decoder use another RNN. The encoder encodes the sequence of text and builds the context vector. Secondly, the decoder read the input sequence from the encoder and generates the predicted output sequence of headline summary.

IV. EXPERIMENTAL SETUP

In this section, explained about experiments that conduct better understands neural headline generation system and describe about data preparation and data processing.

A. Data Preparation

Myanmar news articles are collected written with Myanmar Unicode font from Myanmar News websites. In this experiment, we used a collection of web pages from BBC Burmese news2, Irrawaddy Burmese news3, Weekly Eleven Burmese news4, Mizzima Burmese news5. To build effective headline prediction models, we had to prepare data preparation, data cleaning and the model architecture. Preprocessed data take as the input to the neural network for training and validation. In this data meaningful sentences are included.

TABLE I. DATA STATISTIC OF MYANMAR NEWS DATASET

Data	Article	Sentence
Training Sentences	3211	32110
Validation Sentences	412	4120
Testing Sentences	1377	13770

B. Data Processing

In my training corpus have two column, title and description of text. It can be save CSV format. The titles were used as training target. Before the raw text can be fed into the neural network, it has to be preprocessed. In preprocessing stage sentence segmentation and word segmentation are correctly processed. The input texts are split into an article and summary part. We used preprocessed article text take as input X and the preprocessed summary is target product Y. In data preprocessing includes extracting contents and titles. In this study, training model use the training data 37.7 M news articles with 898 M words. We collected Myanmar news from various Myanmar News websites. The collected news is converted into CSV format and takes as input to the neural architecture for training and validation. Load the CSV input file using python Pandas library [12]. Pandas provide a platform for handling the data in a data frame. It contains many open-source data analysis tool written in Python, the methods to check missing data merge data frames and reshape data structure etc. And then split the dataset into input and output variables for machine learning and apply preprocessing data transforms to the input variables and then finally produce the summarized data as outputs. In training time, we use dimension words embedding by the word2vec algorithm and use epoch for training time. We have used randomly shuffled the training data at every epoch. Training model is developed to identify the sub sequence of words or text that needs to summarize and consider the input text as a sequence of words or text. In extraction phases, summarization model extract the key sentences from the Myanmar news. We implemented our model on the seq2seq and encoder-decoder recurrent networks in Keras library. Keras provides suitable libraries to load the dataset and split it into training set and test set [10]. RNN based Encoder-Decoder model using Google deep learning framework, Tensorflow. Encoder encodes the input Myanmar articles news as the context vector. Decoder reads the encoded input text sequence of text from the encoder and generates the output text sequences.

V. METHOLOGY

A. Neural Headline Generation Model

Neural headline generation targets at learning a model to map a short text into a headline. Neural headline generation means to map documents to headlines with recurrent neural network. Given an input document $X=(x_1,...,x_m)$ where each word x_i get from a fixed vocabulary V, the neural headline generation model to take X as input, and generates a short headline $Y=(y_1,\dots,y_n)$ with length n<m word by word. Headline generator consists of an encoder and decoder which are constructed with Long Short Term

---

[2] https://www.bbc.com/burmese
[3] https://burma.irrawaddy.com/category/news
[4] https://news-eleven.com/news
[5] http://www.mizzimaburmese.com/news
6https://en.wikipedia.org/wiki/Recursive_neural_network

Memory (LSTM), which is one of the recurrent neural networks. The encoder constructs distributed representation from the title sentence of Myanmar article news and decoder generated Myanmar headlines news from the distributed representation.

B. Recursive Recurrent Neural Network

Recursive Recurrent Neural Network (Recursive RNN) is the deep neural network by applying the same set of weight recursively to produce a variable size input structure. Recursive RNN using parse-tree based structure representation.6 Recursive RNN recursively merge pairs of word and phrase representation. Recursive RNN based encoder-decoder neural network architecture. The encoder takes the input text of a news articles one word at a time. Each word is passed through an embedding layer that transforms the words into distributed word representation. The distributed representations of words are combined using multi-layer neural network with the hidden layers generated after serving in the previous word. The decoder must generate two type of information. The first one is context vector (or) fixed length vector and the second one is generated output sequence as a predicted summary. The encoder-decoder architecture means recurrent neural networks for sequence prediction problems that have a variable number of inputs text, outputs text or both inputs and outputs text [11]. The encoder reads the entire input sequence of variable length of text and encodes it into an internal representation of words called the context vector. The decoder reads the encoded the variable input sequence of length from the encoder and generates the output sequence of the predicted headline. In figure 2 shows the encoder and decoder architecture and text is input and headline is generated as output.

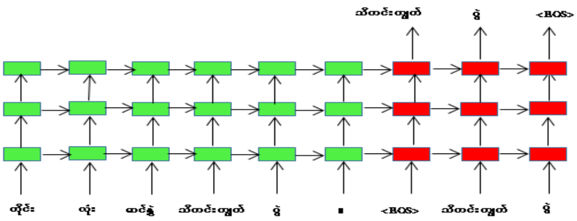

Figure 2. Example Encoder-Decoder neural network architecture

C. Baseline Sequence-to-Sequence with one hot encoding

To verify the effectiveness of the proposed method, has investigated various approaches for comparison. In this work, as the baseline model has considered Sequence to Sequence with one-hot encoding summarization model and Sequence to Sequence with word embedding (GloVe) summarization model. In sequence to sequence one hot encoding model, the sequence provides at least one example of every possible value in the sequence. It uses automatic methods to define the mapping of labels to integers and integers to binary vectors7.

D. Baseline Sequence-to-Sequence with word embedding (GloVe)

Sequence-to-sequence with word embedding (GloVe) summarizer use for encoder input. GloVe means word embedding because the global corpus statistic is captured directly by the model. Global model are widely used for learning distributed word representation. This model is unsupervised learning algorithm for obtaining vector representation for each word GloVe model, as well as some auxiliary tools to construct word-word co-occurrence matrices from large corpora8. The model is an unsupervised learning algorithm for obtaining vector representations for words [14].

VI. EVALUATION

The performance of the model was measured by two different ways. The first way is ROUGE score evaluation and the second way is training and hold out loss. ROUGE (Recall-Oriented Understudy for Gisting Evaluation method based on N-gram statistics, ROUGE-1, ROUGE-2 and ROUGE-L, ROUGE-SU which are computed using the matches of unigrams, bigrams and longest common subsequences and skip-bigram plus unigram-based co-occurrence statistics respectively [13].

---

[7] https://machinelearningmastery.com/how-to-one-hot-encode-sequence-data-in-python/
[8] https://nlp.stanford.edu/projects/glove/

A. Experimental Results

Three models are considered as our assessments evaluation. These models produce the results of original headlines and predicted headlines. In table 2, table 3 and table 4 shows the ROUGE-1, ROUGE-2 and ROUGE-L evaluation results of three models. From the experimental results of Table 2, Table 3 and Table 4, scores are given for ROUGE-1, ROUGE-2 and ROUGE-L, word overlap between a systems generated summary and reference summary. We have noticed that Recursive RNN based summarizer gives the better results of ROUGE-1, ROUGE-2 and ROUGE-L than other two models.

TABLE II. ROUGE-1 SCORE RESULTS FOR THREE MODELS

ROUGE-1			
Model	***Precision***	***Recall***	***F-score***
Seq2Seq with one-hot encoding	0.65377	0.53109	0.56808
Seq2Seq with word embedding	0.36879	0.30019	0.33097
Recursive RNN	**0.71322**	**0.70159**	**0.70736**

TABLE III. ROUGE-2 SCORE RESULTS FOR THREE MODELS

ROUGE-2			
Model	***Precision***	***Recall***	***F-score***
Seq2Seq with one-hot encoding	0.28497	0.22821	0.25345
Seq2Seq with word embedding	0.11774	0.08671	0.09609
Recursive RNN	**0.37352**	**0.36706**	**0.37027**

TABLE IV. ROUGE-L SCORE RESULTS FOR THREE MODELS

ROUGE-L			
Model	***Precision***	***Recall***	***F-score***
Seq2Seq with one-hot encoding	0.12234	0.13241	0.23385
Seq2Seq with word embedding	0.09052	0.09987	0.14902
Recursive RNN	**0.11222**	**0.18937**	**0.31844**

Table 5 shows the example of generated headlines for articles from various Myanmar news websites.

In figure 3, figure 4 and figure 5 shows the training plot below was obtain by running for 100 epochs by using Recursive RNN model, Sequence to Sequence with one hot encoding model and seq2seq with word embedding model. Line plot shows the y-axis and x-axis. Y-axis shows the accuracy of the model and x-axis shows the number of epoch. In figure 3 shows the evaluation metrics as a function of training epoch and we can see that the train accuracy values are relatively higher than the validation accuracy. In figure 3 shows the model in which training loss suddenly decreases but validation loss eventually goes up.

Accuracy vs Loss (Recursive RNN)

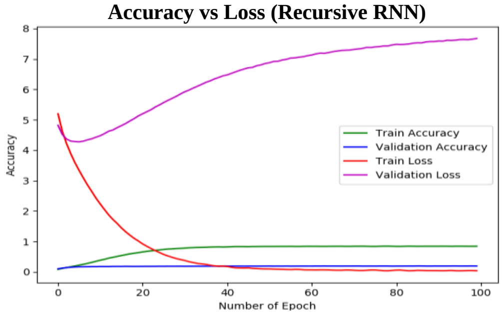

Figure 3. accuracy vs loss with Recursive RNN

In Fig. 4 shows the learning curve of Seq2Seq with one hot encoding model. From the output, we can see that the train accuracy slightly higher than the validation accuracy.

Accuracy vs vs Loss (Seq2Seq with one hot encoding)

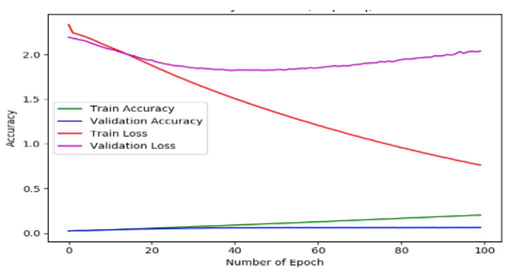

Figure 4. accuracy vs loss with seq2seq with one hot encoding

In fig 5 shows the learning curve of Seq2Seq with word embedding (GloVe) model. The output result shows that difference between train accuracy and validation accuracy. According to the training results, Recursive RNN-based headline summarization model is good performance results than other two models.

TABLE V. EXAMPLE OF MYANMAR NEWS PREDICTED HEADLINE

Text	Actual Headline	Predicted Headline
အောင်လအန်ဆန် ၏ ပွဲစဉ် အား ကြည့်ရှုအားပေး စဉ် အသက် ၇၀ ကျော် အဘိုးအို တစ်ဦး လဲ ကျ သေဆုံးမှု ဒေးဒရဲ မြို့ တွင် ဖြစ်ပွား ခဲ့ ကြောင်း သိရသည် ။ ဧရာဝတီ တိုင်းဒေသကြီး ဒေးဒရဲ မြို့ မော်ဘိ စု ကျေးရွာအုပ်စု ထ နိ ပတ် ကျေးရွာ တွင် အောက်တိုဘာ ၁၃ ရက် ညပိုင်း တွင် “ မြန်မာ့ ဂုဏ် ဆောင် အောင်လ ” ဟု အော်ဟစ် အားပေး ကာ လဲ ကျ သေဆုံး ခဲ့ ခြင်း ဖြစ်ကြောင်း ကျေးရွာ အုပ်ချုပ် ရေး မှူး ရုံး မှ သိရသည် ။	အောင်လအန်ဆန် ၏ ပွဲစဉ် အား ကြည့်ရှု အားပေး စဉ် အသက် ၇၀ ကျော် အဘိုးအို တစ်ဦး လဲ ကျ သေဆုံး မှု ဒေးဒရဲ မြို့ တွင် ဖြစ်ပွား	အောင်လအန်ဆန် ၏ ပွဲစဉ် ကြည့်ရှု အားပေး စဉ် အသက် ၇၀ ကျော် အဘိုးအို တစ်ဦး သေဆုံး မှု ဖြစ်ပွား
၆၆ နှစ်မြောက် တရုတ် ပြည်သူ သမ္မတနိုင်ငံ ထူထောင်ခြင်း အခမ်းအနား ကို မြန်မာနိုင်ငံ ဆိုင်ရာ တရုတ် သံရုံး မှ ပြည်ထောင်စု နယ်မြေ နေပြည်တော် တွင် ပထမဆုံး အကြိမ် အဖြစ် စက်တင်ဘာ ၂၄ ရက် တွင် ကျင်းပခဲ့သည် ။ မြန်မာ အပြည်ပြည်ဆိုင်ရာ ကွန်ဗင်းရှင်း စင် တာ (MICC) ၌ ကျင်းပ သည့် အခမ်းအနား တွင် ၂၀၁၅ ခုနှစ် သည် တရုတ် – မြန်မာ သံတမန် ဆက်ဆံရေး တည်ဆောက် ခဲ့ သည့် ၆၅ နှစ် ပြည့် သည့် နှစ် လည်း ဖြစ်သည် ဟု မြန်မာနိုင်ငံ ဆိုင်ရာ တရုတ် သံအမတ်ကြီး မစ္စတာ ဟုန် လျန် က ပြောသည် ။ တရုတ် သံအမတ်ကြီး မစ္စတာ ဟုန်လျန် က ၂၀၁၁ ခုနှစ် တွင် တရုတ် – မြန်မာ နှစ်နိုင်ငံ ဘက်စုံ မဟာဗျူဟာ မိတ်ဆွေ ဆက်ဆံရေး တည်ထောင် နိုင် ခဲ့ ပြီး ၊ တရုတ် – မြန်မာ နှစ်နိုင်ငံ သည် အနီးကပ် မိတ်ဆွေ အိမ်နီးချင်း နိုင်ငံ အဖြစ် ကံကြမ္မာ အကျိုးတူ အသိုက်အဝန်း ကိုလည်း ဖြစ်ပေါ် စေ ခဲ့ သည် ဟု တရုတ် သံအမတ်ကြီး ပြောသည် ။	၆၆ နှစ်မြောက် တရုတ် ပြည်သူ သမ္မတ နိုင်ငံ ထူထောင်ခြင်း ခမ်းအနား ကို မြန်မာနိုင်ငံ ဆိုင်ရာ တရုတ် သံရုံး မှ ပြည်ထောင်စု နယ်မြေ နေပြည်တော် တွင် ပထမဆုံး အကြိမ် အဖြစ် စက်တင်ဘာ ၂၄ ရက် တွင် ကျင်းပခဲ့	၆၆ နှစ်မြောက် တရုတ် ပြည်သူ သမ္မတ နိုင်ငံ ထူထောင်ခြင်း အခမ်းအနား ကို မြန်မာ နိုင်ငံ ဆိုင်ရာ တရုတ် သံရုံး မှ ပထမဆုံး အကြိမ် အဖြစ် ကျင်းပခဲ့
"လာ မည့် နိုဝင်ဘာလ ၈ ရက် အထွေထွေရွေးကောက်ပွဲ အတွက် ထိုင်းနိုင်ငံ ဘန်ကောက် မြို့ နှင့် ချင်းမိုင် မြို့ တို့တွင် ကြိုတင်မဲပေး ရန် လာရောက် လျှောက်ထားသူ စုစုပေါင်း ၈၅၄ ဦး မှ ၆၀၄ ဦး သာ မဲပေး ခွင့် ရှိပြီး ကျန် ၂၅၀ မှာ မဲ စာရင်း တွင် မ ပါဝင် ကြ ဟု သိရသည် ။ ဘန်ကောက် မြို့ မြန်မာ သံရုံး က ပြီး ခဲ့ သည့် ဇူလိုင်လ နှင့် စက်တင်ဘာလ များ တွင် ကြိုတင်မဲပေး ရန် အသိပေး ခဲ့ ပြီးနောက် လာရောက် လျှောက်ထားသူ ပေါင်း ၇၁၂ ဦး ရှိရာ မှ ၅၆၁ ဦး သာ မဲပေး ခွင့် ရပြီး ၊ မဲ စာရင်း တွင် မပါ သူ ၁၅၁ ဦး ရှိသည် ဟု ထိုင်းနိုင်ငံ ဘန်ကောက် မြို့ မြန်မာ သံရုံး မှ သံအမတ်ကြီး ဦး ဝဏ္ဏဟန် က အောက်တိုဘာလ ၁၅ ရက် ၌ မဇ္ဈိမ ကို ပြောသည် ။	လာ မည့် နိုဝင်ဘာလ ၈ ရက် အထွေထွေ ရွေးကောက်ပွဲ အတွက် ထိုင်း နိုင်ငံ ဘန်ကောက် မြို့ နှင့် ချင်းမိုင် မြို့ တို့တွင် ကြိုတင်မဲပေး ရန် လာရောက် လျှောက်ထားသူ စုစုပေါင်း ၈၅၄ ဦး မှ ၆၀၄ ဦး သာ မဲပေး ခွင့် ရှိ	လာ မည့် နိုဝင်ဘာလ ၈ ရက် အထွေထွေ ရွေးကောက်ပွဲ အတွက် လာရောက် လျှောက်ထားသူ စုစုပေါင်း ၈၅၄ ဦး မှ ၆၀၄ ဦး သာ မဲပေး ခွင့် ရှိ

Accuracy vs loss (Seq2Seq with word embedding GloVe)

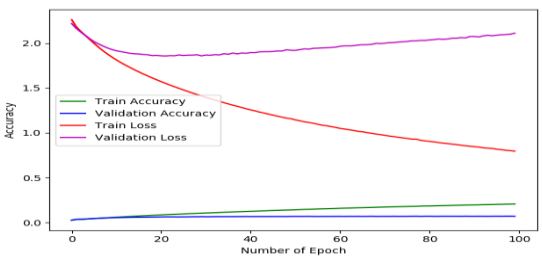

Figure 5. accuracy vs loss with Seq2Seq with word embedding (GloVe)

VII. CONCLUSION AND FUTURE WORK

We trained an encoder-decoder recurrent neural network model for generating Myanmar news headlines using the text of Myanmar news articles. The model generates a concise summary as generated outputs. Recursive RNN can be more learn complex linguistic large amount of training data. Moreover evaluations on experimental summarized results have been also discussed with related evaluation scores. According to the evaluation results, this experiment can easily provide predicted headline summaries from Myanmar news web page with acceptable evaluation scores within a particular time as the goal of our motivation. In my future work, by using larger training data sets and training these systems as auto encoders, where input is expected to match output.

References

[1] “Abstractive text summarization using sequence-to-sequence RNN and beyond", Ramesh Nallapati, Bowen Zhou, Caglar Gulcehre, Bing Xiang, 2016/2/19

[2] "Abstractive and Extractive Text Summarization use Document Context Vector and Recurrent Neural Networks." arXiv preprint arXiv: 1807.08000 (2018). Khatri, Chandra, Gyanit Singh, and Nish Parikh

[3] “Neural abstractive summarization” was pioneered in Rush et al. (2015),

[4] "Ranking with recursive neural networks and its application to multi-document summarization" Cao, Ziqiang,et al Twenty-ninth AAAI conferences on artificial intelligence. 2015.

[5] "Abstractive Text Summarization Using Artificial Intelligence”.Parmar, Chandu, Ranjan Chaubey, and Kirtan Bhatt Available at SSRN 3370795 (2019).

[6] "Sequence to sequence learning with neural networks." Sutskever, Ilya, Oriol Vinyals, and Quoc V Advances in neural information processing systems. 2014.

[7] "Abstractive Summarization for Amazon Reviews." Yang, Lu. (2016).

[8] “Myanmar word segmentation using hybrid approach”. In Proc. of ICCA. 166–170. Win Pa Pa and Ni Lar Thein. 2008. [9] "Fast abstractive summarization with reinforce-selected sentence rewriting." Chen, Yen-Chun, and Mohit Bansal. arXiv preprint arXiv:1805.11080 (2018).

[9] "Fast abstractive summarization with reinforce-selected sentence rewriting." Chen, Yen-Chun, and Mohit Bansal. arXiv preprint arXiv:1805.11080 (2018).

[10] “Deep Learning with Keras. Gulli, Antonio, and Sujit Pal. Packt Publishing Ltd, 2017.

[11] "Machine Learning Approach for Automatic Text Summarization Using Neural Networks‖." Patel, Meetkumar, et al. International Journal of Advanced Research in Computer and Communication Engineering 7.1 (2018).

[12] " Python for High Performance and Scientific Computing 14 “,McKinney, Wes. "pandas: a foundational Python library for data analysis and statistics (2011).

[13] "Rouge: A package for automatic evaluation of summaries." Lin,Chin-Yew. Text Summarization Branches out (2004).

[14] Glove: Global vectors for word representation." Pennington, Jeffrey, Richard Socher, and Christopher Manning Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). 2014.