Myanmar Dialogue Act Recognition (MDAR)

Sann Su Su Yee
Natural Language Processing Lab
University of Computer Studies, Yangon
Myanmar
sannsusuyee@ucsy.edu.mm

Khin Mar Soe
Natural Language Processing Lab
University of Computer Studies,Yangon
Myanmar
khinmarsoe@ucsy.edu.mm

Ye Kyaw Thu
National Electronics and Computer Technology Center (NICT)
Thailand
yktnlp@gmail.com

Abstract

This research aim to make the very first machine learning based Myanmar Dialog Act Recognition (MDAR) for Myanmar Dialogue System. As we know, Dialog Act (DA) recognition is the early level of dialogue understanding which can capture aspects of the user, and they are sentence-level units that represent states of a dialogue, such as greeting, question, inform, and so on. We focus on the current works about DA recognition, especially for Myanmar Dialogue. In this work, we used two machine learning approaches, which are Naïve Bayes classifier and Support Vector Machine (SVM), for dialogue act tagging in the MmTravel (Myanmar Travel) corpus, and the results of two approaches are slightly different but the result of SVM approach attained in the term of average F-measure scores of 0.79; showed that these approach has moderately good accuracy for Myanmar dialogue.

Keywords—*Myanmar Dialogue Act Recognition (MDAR), Naïve Bayes, Support Vector Machine (SVM), MmTravel corpus*

I. INTRODUCTION

Natural Language Understanding (NLU) is an important component of dialogue management, and NLU has been extraordinarily improved by deep learning techniques, but current NLP techniques for Myanmar language dialogue action classification is new research area where it has been difficult to infer a dialogue act from a surface utterance because it depends on the context of the utterance and speaker linguistic knowledge. From linguistic perspectives on NLU, Allen [15] describes the following forms of knowledge which are phonetic and phonological knowledge, morphological knowledge, syntactic knowledge, semantic knowledge and pragmatic knowledge. Among them, pragmatic knowledge which is focused on how sentences are used in different situations and how to make interpretations of the sentences. In there, one of the subfields of pragmatics, speech act, that studies how words are used not only to present information but also to carry out actions. Dialogue act is also a type of speech act.

Dialogue act recognition is an essential task for dialogue systems. Automatically Dialogue Act Modeling [13][3] and detecting the structure of dialog is critical to better interpret and help in understanding a conversation. The minimal units of linguistic communication of DAs are directly connected with the speaker's communicative intentions. Austin defines the dialogue act is the function of a sentence in the dialogue which means the function of a question is to request some information, while an answer shall provide this information. The former state of act recognition has been addressed by Searle in 1969, which is based on Austin (1962) work as a fundamental concept of linguistic pragmatics, analyzing, for example, what it means to ask a question or make a statement. To the best of our knowledge, there is no research work on Myanmar dialogue act recognition and modeling that have been published in NLP, but there are several works for other languages, especially for English. Different sets of dialogue acts are defined depending on the target application. In [4], standard of discourse structure annotation, the Dialog Act Markup in Several Layers (DAMSL) tag set [14], which designed by the natural-language processing community (Core & Allen) in 1997. In total, 42 dialogue act classes were defined for English. Switchboard-DAMSL [6] tagset is the modified of DAMSL in the telephone conversation domain. And, another popular tagset is the Meeting Recorder Dialogue Act (MRDA) tagset [MRDA], where contains 11 general DA labels and 39 specific labels, which is based on the taxonomy of SWBD-DAMSL. The Map-Task [12] is English tagset, which contains 19 tags. AMI [11], DIT++ [9] and other ISO standards [10] are examples of dialogue act tagset.

Many researchers have been widely explored over the years for dialogue act recognition as a task, using multiple classical machine learning approaches, Hidden Markov Models (HMM) [17], Maximum Entropy (Maxent), Bayesian Network, and Support Vector Machines (SVMs) [5][8]. Recently, the research of DA recognition approaches applies by Neural Network which will be our future work for Myanmar dialogue. The authors proposed a fully-automated method for the task of speech act classification for Arabic discourse, as in [1]. Arabic sentences have been collected from two Arabic news sources: Al-Hayat newspaper and Aljazeera television station, and defined 10 Arabic speech act. Naïve Bayes and Decision Trees algorithms were used to induce speech act classifiers for Arabic texts and used as features by the word-tag pairs. Decision-tree based classifiers worked slightly better for a 3-word context while Naïve Bayes were better for 4- or 5-word contexts. Results obtained using sequences (bigrams and trigrams) of POS tags had higher accuracy when using naïve Bayes classifier but lower or similar accuracy scores when using decision trees. For model evaluation, they evaluated by 10-fold cross-validation. Reference [2] contributed a combination of a Naïve Bayes classifier and n-grams based natural speech dialogue act detection on two corpora: the Switchboard and the Basurde tasks. 66% of accuracy is achieved on the Switchboard corpus by using a uniform Naïve Bayes classifier. And also, it has been used 3-grams and Laplace smoothing to avoid zero probabilities. For the Basurde corpus, they applied Naïve Bayes classifier with 2-grams and Written Bell smoothing and achieved the best accuracy of 89%.

For multiclass DA classification, [16] applied SVMs on the ICSI meeting corpus that consists of 75 natural occurring meetings and each meeting has long an hour by five participants. Act tags are grouped into 5 broad intuitive classes. In their evaluation, they mention combining multiple binary SVMs via error correction output codes (ECOC) achieved better performance than representing a direct multiclass SVM. Reference [7] used the SVM linear kernel combine with HMM models in the HCRC MapTask corpus which is a collection of 128 2-speaker dialogs with twelve tags. The corpus's each dialog has the direction of a shared map from speaker to listener. The authors approved the SVM can easily integrate sparse high-dimensional text features and dense low-dimensional acoustic features. The classification accuracy of 42.5% and 59.1% respectively for acoustic and text features with a linear SVM followed by Viterbi decoding, and 65.5% for combination.

II. MMTRAVEL CORPUS

We proposed Myanmar dialogue corpus which is based on ASEAN MT[16] Myanmar dataset and online resources for travel domain. In our corpus, we also modified ASEAN MT data to an informal conversation between humans, and it has 60k utterances about question and answer conversation. The corpus has longer utterance, on average, there are 30 words and short utterance consists 5 words. The first step in establishing a dialogue act recognition system is defining the appropriate functions or the DA tag-set. Different types of dialogue systems require labeling different kinds of acts. Therefore, we categorized twenty-nine DA tagsets based on five kinds of Myanmar speech function is listed in Table 1 [18], [19], [20], [21], [22], [23], [24].

A. Informational Function

Informational function concentrates on the message between speaker and listener. It is used to give new information and it depends on truth and value. Inform (inf) act is included in this function. It conveys information to make someone aware of something and listener can decide that information is right or wrong. Example: "အကောင်းဆုံး ကော်ဖီမှုန့် သုံးထားတာပါ (the best coffee powder are used)".

B. Expressive Function

The expressive function can be used to express attitudes and feelings. Accept (ac) describes used to agree to take something and to say 'yes' to an offer or invitation. An act of saying sorry is the apology (apol) act. Congratulate (cong) describes when we praise (someone) for an achievement, for example: "ဝိုး တကယ် လား ၊ ဂုဏ်ယူ ပါတယ် (Wow really, congratulation)". Complain act (cp) expresses dissatisfaction or annoyance about something. Deny (dny) act used to not allow someone to have or do something and state that one refuses to admit the truth or existence of, example sentence is "ခင်ဗျား ပြောတာ မဖြစ်နိုင်ဘူး (It's impossible)". Opinion (op) act is a thought or belief about something or someone, like this sentence "ပန်းသီး ဆိုရင် ကောင်းမယ် (Apple is much better)". Our corpus is based on traveling dialogue therefore the shopping and eating is a kind of categories where review (rev) act to evaluate a service or food. Wish (wn) describes use to hope or express hope for another person’s success or happiness or pleasure on an occasion, for example “လမ်းခရီး တစ်လျှောက်လုံး အဆင်ပြေပါစေ (Have a safe journey)”**.**

C. Directive Function

Directive function, which aims to influence the behavior or attitudes of others. There are many kinds of act tag in here:

- Command (cmd) describes acts used to control over someone or something and tell them what to do, for example “မလုပ်နဲ့ (Don’t do it)”**.**
- Direction (dir) used to instruct someone about how to find a particular place and pointing (someone) towards.
- Invite (inv) used to invite or request someone to come or go to some places, especially formally or politely.

TABLE I. PROPOSED MYANMAR DIALOGUE ACT TAGSETS

Speech Function	Dialogue Act	Abbrev
Informational	Inform	inf
Expressive	Accept/Agree	ac
	Apology	apol
	Congratulate	cong
	Complain	cp
	Deny	dny
	Opinion	op
	Review	rev
	Wish	w
Directive	Command	cmd
	Directions	dir
	Invite	inv
	Instruction	instr
	Urge	u
	Prohibit	proh
	Request	req
	Suggestion	sug
	Thank	thx
	Warning	wn
	Confirm Question	cfm_q
	Choice Question	ch_q
	Complain Question	cp_q
	Inquiry Question	inq_q
	Request Question	req_q
	Other Question	otr_q
Phatic	Greeting	gt
	Goodbye	gb
	Self_Intro	s_i
Aesthetic	Aesthetic	as

- Instruction (instr) used to advice and information about how to do or use something.
- Urge (u) used to strongly advise or try to persuade someone to do a thing.
- Prohibit (proh) used to officially refuse to allow something.
- Request (req) which is the act of politely or officially asking for something.
- Suggestion (sug) includes an idea, plan, or action that is suggested or the act of suggestion it.
- Thank (thx) describes acts used to express to someone that are grateful for something that they have done.
- Warning (wn) is something that makes you understand there is a possible danger or problem, especially one in the future.
- Question has (6) kind of act tag: Confirm_Q (cfm_q), Choice_Q (ch_q), Complain_Q((cp_q), Inquiry_Q (inq_q), Request_Q (req_q), and Other_Q (otr_q), where we tagged them in sentence or phrase used to find out information.

D.Phatic and Aesthetic Function

Phatic function uses to maintain social relationships, and to begin or continue the conversation, which helps us to interact with people. Greeting (gt), goodbye (gb), and self_intro (s_i) are kind of phatic function. Example: “မင်္ဂလာ နေ့လည်ခင်းပါ ၊ တွေ့ရတာ ဝမ်းသာပါတယ် (Good afternoon, glad to meet you)”**.** The aesthetic function includes the rhythm, balance and contrast of sentence, clauses and words also play.

TABLE II. PERCENTAGE OF UTTERANCES ASSIGNED TO EACH DIALOG ACT CONSIDERED IN THE CORPUS

No.	act_tag	%	Example
1	inf	35.70	ဘာ လို့ လဲ ဆိုရင် ဒီမှာ လတ်ဆတ်တဲ့ အသီးတွေ ကျွန်တော် တွေ့ခဲ့ ပါတယ်
2	inq_q	31.78	ခင်ဗျား လမ်းညွှန်မြေပုံ ပါ လား
3	req	6.90	နောက် တစ်ပတ် သောကြာနေ့ မတိုင်ခင် မန္တလေး မြို့ ကို သွား မယ့် လေယာဉ် မှာ လက်မှတ် တစ်စောင် ကြိုတင် မှာ ချင်လို့ပါ
4	op	5.22	အကောင်းဆုံး လက်ဝတ် ရတနာ ကိုး
5	ac	4.07	အင်း ယူ သွား ပေး မယ်
6	req_q	3.30	ရှင် ကျေးဇူးပြုပြီး ဒါ ကို ပင်မင်းဆိုင် ယူ သွား ပေး မလား

No.	act_tag	%	Example
7	instr	1.40	အဲဒါတွေ ကို ကား ပေါ် တင်လိုက် ပါ
8	thx	1.36	စောင့်ပေး လို့ ကျေးဇူးတင် ပါတယ်
9	apol	1.03	ကျွန်တော် နောက်ကျသွား လို့ တောင်းပန် ပါတယ်
10	dny	1.00	မင်း ကို အဝတ်အစား ဝယ် ဖို့ ငါ တော့ ပိုက်ဆံ တစ်ပြား မှ မ ထုတ်ပေး နိုင် ဘူး
11	sug	1.00	လေဆိပ် ကို မနက် ၈ နာရီ မတိုင်ခင် ရောက်ရင် ကောင်း ပါတယ်
12	gt	1.00	မင်္ဂလာ ပါ ၊ နေကောင်း လား
13	rev	0.84	ကျွန်တော် သိ သလောက်တော့ ဒါဟာ ရှိသမျှ ထဲမှာ တကယ့် အကောင်းဆုံး ပဲ
14	u	0.84	အကောင်းဆုံး ကို မျှော်လင့် ကြ ပါ စို့
15	dir	0.71	ဟိုဘက် လမ်း မှာ ဈေးဝယ်စင်တာ ရှိတယ်
16	cp	0.64	မင်း ကြောင့် ခုတော့ ငါ့ မှာ နောက်ထပ် အထုပ် တစ်ထုပ် ထပ်ပို့ ရဦး မှာ ပေါ့
17	cfm_q	0.57	ခင်ဗျား ကျွန်တော့် ကို တစ်ခုခု ယူလာ စေချင် တာ လား
18	ch_q	0.55	အအေး လား အပူ လား
19	proh	0.45	ဒီနေရာ က ဆေးလိပ် ကင်းစင် နေရာ မို့ ဆေးလိပ် မသောက်ပါနဲ့
20	w	0.43	မင်း ရဲ့ လေကြောင်းခရီး မှာ ပျော်ရွှင် ပါစေ
21	cmd	0.37	တိတ်တိတ် နေ
22	s_i	0.19	ကျွန်မ နာမည် သူဇာ ပါ
23	as	0.18	အလျင်လို အနှေးဖြစ် ဆိုတာလိုပဲ
24	gb	0.15	ကျွန်တော့် ကို ခွင့်ပြု ပါဦး မနက်ဖြန် တွေ့ ပါမယ်
25	inv	0.12	ခင်ဗျား ကို ကျွန်တော်တို့ မိသားစု က အမြဲ ကြိုဆို ပါတယ်
26	wn	< 0.1	နောက် ကို ကျွန်တော် ပြော တဲ့ အတိုင်း လုပ်ပါ ၊ ဒါ နောက်ဆုံး အကြိမ် ဖြစ် ပါစေ
27	cp_q	< 0.1	ဘာ လို့ အခုထိ ရေ မလာသေးတာ လဲ ၊ အဆင်ကို မပြေဘူး
28	cong	< 0.1	ဟုတ် လား ၊ ဂုဏ်ပြု ပါတယ် နော်
29	otr_q	< 0.1	မင်း ဘာ ကို မ ဝယ် ခဲ့ သလဲ ဆိုတာ

Aesthetic (as) tag example sentence is "လူပြောမသန် လူသန်မပြော".

The most top five frequent DA types include INFORM, INQUIRY QUESTION, REQUEST, OPINION, and AGREEMENT/ACCEPT. We list of all dialogue acts in our corpus ordered by the highest frequencies together with the example sentence can be seen in Table 2, where Myanmar sentences translation are given in Appendix A. Generally, our daily conversation tends to give information about something and implies the imparting of knowledge especially of facts or occurrences to the listener, that has been made the INFORM (inf) act tag to increase the percentage.

III. METHODOLOGY

Dialogue act classification is a special case of text classification, where the samples are the utterances, and the attributes are the words of that.

A. Naïve Bayes Classifier

The supervised learning method, Bayesian Classification, is also a statistical method to classify document or text. Its algorithm based on applying Bayes theorem with naïve assumption which means every feature is independent of the others, in order to predict the category of a given sample. The probabilistic model calculates the probability of each category using Bayes theorem, and categorize with the highest probability will be output. To find a function f*(.) which maps an utterance u_j into an act label a_j. The range is defined by the number of act labels A, where includes 29 dialogue acts of the MmTravel task. In the training, the learning function is done from a set of samples with the form:

$$\{(u_j, a_j)\}_{J=1}^{J},\ u_j \in U,\ a_j \in A \qquad (1)$$

where u_j is the j-th sample, J is the number of samples, and a_j correspond act label. The decision of act tag is assigned to each utterance, which is made by the Bayes decision rule for minimizing error. The classifier assigns the act with maximum probability to the utterance u:

$$a_{MAP} = \underset{a \in A}{\operatorname{argmax}}\ P(a|u) \qquad (2)$$

where MAP is "maximum posteriori" which means most likely class are assigned the result. By Bayes rule,

$$= \underset{a \in A}{\operatorname{argmax}} \frac{P(u|a)P(a)}{P(u)} \qquad (3)$$

For dropping the denominator,

$$= \underset{a \in A}{\operatorname{argmax}}\ P(u|a)P(a) \qquad (4)$$

$$a_{MAP} = \underset{a \in A}{\operatorname{argmax}}\ P(x_1, x_2, ..., x_n|a)P(a) \qquad (5)$$

where utterance u represented as features $x_1, x_2, \ldots, x_n$.

B. Support Vector Machine

SVM is one of the robust classification models with good generalization ability on unseen data, and it is also a binary classifier where it works on the concept of identifying the separating hyper plane, as decision boundary, with maximum margin between two classes. The objective of SVM is to define the decision boundary with the maximum margin, where it can be defined the distance from the nearest training example to the decision boundary. SVM classifier can be used to classify not only two given classes but also multiple classes.

The basic classifier is linear classifier which separate a linearly separable data. In non-linear decision boundary SVM draw boundary by transforming the input original space to a high dimensional space. The function transforming data to a non-linear space is called Kernel functions. If the classifier can determine the optimal separating hyper plane, it guarantees to give better result for its model. The decision boundary is determined by the discriminant:

$$f(x) = \sum_i a_i \lambda_i K(u,u_i) + b \qquad (6)$$

where u_i and $a_i \in \{-1,1\}$ are the utterance-tag pairs for training set $\{(u_i, a_i)\}_{i=1}^{N}$. The kernel function, $K(u, a) \doteq \phi(u) \cdot \phi(a)$, which computes inter products, and $\phi(u)$ is a transformation from the input space to a higher dimensional space. $\phi(u) = u$ used for the linearly separable case. In non-linear separable cases, an SVM apply the mapping $\phi(.)$ to increase dimensionality and after that applying a linear classifier in the higher dimension. Kernel functions $K(.,.)$ of SVM classifier are listed below:

For Linear Kernel,

$$K(u,u_i) = u_i^T u \qquad (7)$$

For Polynomial Kernel,

$$K(u,u_i) = (u_i^T u + \tau)^d \qquad (8)$$

For Radial Basis Function (RBF) Kernel,

$$K(u,u_i) = \exp(-\| u - u_i \|_2^2 / \sigma^2) \qquad (9)$$

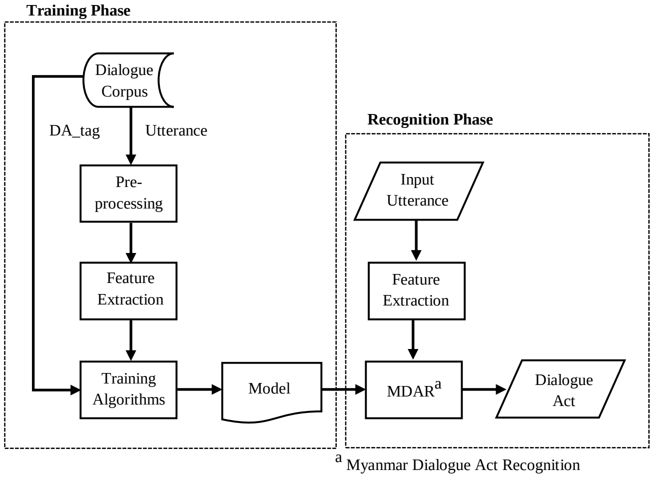

Figure 1. Flow chart of the proposed Myanmar Dialogue Act Recognition

IV. MYANMAR DIALOGUE ACT RECOGNITION

Dialogue act recognition task is the same as classification which acts a dialogue label to sentence, where multiple classification methods have been used to do this undertaking. Most of them were supervised approaches, which means that acquired a large amount of annotated data to obtain models. For this reason, we firstly prepared MmTravel corpus which coverage 12 categories, where includes transportation, tourism, sightseeing, shopping, healthcare, food, emergency, communication, airport, accommodation and other online resources for travel domain. The illustration of MDR proposed flow chart is shown in Fig. 1.

A. Pre-processing

In pre-processing phase, we applied Myanmar word segmentation tool[1] that published by UCSY[2] NLP Lab. We split the dataset by using cross-validation with 80% for the training set and 20% in the test set to perform the experiments.

B. Feature Extraction

Feature extraction used to extract features from datasets to a format supported by machine learning approaches. DA classification is very similar to sequence classification problem. Most of the other languages dialogue act detection system uses the feature-based statistical classification combine with several feature such as linguistic and prosody, by using appropriate machine learning approaches. But in our work, we use tfidfvectorizer (term frequency-inverse document frequency vectorizer) which turning a

[1] http://www.nlpresearch-ucsy.edu.mm/NLP_UCSY/wsandpos.html
[2] University of Computer Studies, Yangon – www.ucsy.edu.mm

collection of raw data into numerical feature vectors and show the resulting scores assigned to each word. Term frequency summarizes how often a given word appears within a document, and inverse document frequency downscales words which appear a lot across documents. The statistical measure, tf-idf weight: $w_{i,j} = tf_{i,j} \times \log(\frac{N}{df_i})$, used to evaluate which word is important in an utterance in a corpus.

C. Training

For the training algorithm, we analyzed MultinomialNB and BernoulliNB for Naïve Bayes classifier, and different kernel-based SVM machine learning on our annotated data. Some of the top features extracted by MultinomialNB and BernoulliNB classifier on our datasets are listed in Table 3 and 4 respectively. Among them, some of the common words in

TABLE III. TOP FEATURES BY MULTILNOMIAL NAÏVE BAYES

Dialogue Act		Features Words
Request Question	req_q	ဝမ်းနည်း ၊ နိုင်မလား ၊ လုပ်ပေး ၊ ပေးပါ
Directions	dir	ဒီကား ၊ ဒီလမ်း ၊ ဒီမှာပဲ ၊ ဒီည ၊ ဒီညနေ ၊ ဒီထက် ၊ ဒီနား ၊ ဒီနေရာ ၊ ဒီဘက်
Command	cmd	ထည့် ၊ ဖို့ ၊ ရဘူး ၊ နော် ၊ တော့ ၊ ကို ၊ မ ၊ နဲ့
Greeting	gt	ဟဲလို ၊ မင်္ဂလာ ၊ **ကျွန်တော်** ၊ ဟိုင်း ၊ ဝမ်းသာ ၊ ရဲ့ ၊ နေကောင်း ၊ **ခင်ဗျား**

TABLE IV. TOP FEATURES BY BERNOULLI NAÏVE BAYES

Dialogue Act		Features Words
Instruction	instr	ရမယ် ၊ သောက် ၊ ပါ ၊ **မင်း** ၊ က ၊ ဖို့ ၊ ဒီ ၊ နဲ့ ၊ ကို
Apology	apol	အားနာ ၊ စိတ်မကောင်းပါဘူး ၊ တောင်းပန်
Deny	dny	မဟုတ်ပါဘူး ၊ ဘူး ၊ မဟုတ်ဘူး ၊ မ ၊ ဟင့်အင်း ၊ နဲ့
Wish	w	မျှော်လင့် ၊ ပျော်ရွှင် ၊ မှာ ၊ ခရီး ၊ **ကျွန်တော်** ၊ ပါစေ

Myanmar language as ကျွန်တော် (“I”), မင်း (“You”), ခင်ဗျား (“You”), are appear frequently in the corpus. So, we will be estimated classification by deleting stopwords from the utterances in our future work.

Support Vector Machine (SVM), as a powerful supervised learning method, is suitable to deal with classification problems and has high accuracy rate for linearly inseparable data, which can be mapped into a high dimensional space through kernel functions. In our experiment, we study the impact of kernel functions which are linear kernel, RBF kernel, and Polynomial kernel. Our experiment among these three kernels, the learning ability of RBF kernel is stronger than the other kernels, we also express the training and evaluation report for RBF kernel as follows:

Accuracy on Training Set : 0.8568779
Accuracy on Testing Set : 0.7913935
နောက်ထပ် ရော ဘာ ဖြစ်သေး လဲ
Predicted Target: inq_q
Actual Target: inq_q
မင်း အချိန် ရှိလာ မှာ ပါ
Predicted Target: inf
Actual Target: op
ဒါလေး ယူသွား ပေးပါ လား
Predicted Target: req_q
Actual Target: req_q
ခင်ဗျား စီးချင် တဲ့ လေကြောင်းလိုင်း ကို ခေါ် ပြီး
ခင်ဗျား ဘာတွေ သယ်နိုင် လဲ ဆိုတာ မေးနိုင် ပါတယ်
Predicted Target: inq_q
Actual Target: u
အခုလို တွေ့ ရတာ ဝမ်းသာ ပါတယ်
Predicted Target: gt
Actual Target: gt
ဒီလမ်း အတိုင်း တည့်တည့် လျှောက် ပါ ပြီးရင် ညာကွေ့
Predicted Target: dir
Actual Target: dir

[3]Example sentences translation are given in Appendix B

Table 5. shows the classification report for (29) dialogue act tags by using Precision (P), Recall (R), and the F1 score (F1). But the result for act tag: as, cong, gb, inv, and otr_q, are less than 0.1% because of the small number of these tags contained in our corpus, and we will be considered the data balancing for all tagsets in our future work.

D. Experiment

We presented the classification result of the Naïve Bayes and each kernel type of SVM classifier for the DA classification task is shown in Table 6. Overall, from the range of classification undertaken, SVM RBF kernel was obtained the highest average classification accuracy at 79 % compares favorably to the other classifiers and it performs slightly better than linear and polynomial kernels, suggesting that the RBF kernel approach may be able to capture some discriminative features that directly distinguish different classes.

TABLE V. CLASSIFICATION RESULT OF RBF KERNEL

Dialogue Act		Precision	Recall	F1
Accept/Agree	ac	0.73	0.65	0.68
Apology	apol	0.90	0.77	0.83
Aesthetic	as	< 0.1	< 0.1	< 0.1
Confirm Question	cfm_q	0.53	0.15	0.24
Choice Question	ch_q	0.72	0.34	0.46
Command	cmd	0.83	0.14	0.24
Congratulate	cong	< 0.1	< 0.1	< 0.1
Complain	cp	1.00	0.08	0.14
Complain Question	cp_q	0.50	0.20	0.29
Directions	dir	0.78	0.27	0.40
Deny	dny	0.84	0.47	0.60
Goodbye	gb	< 0.1	< 0.1	< 0.1
Greeting	gt	0.87	0.58	0.69
Inform	inf	0.72	0.93	0.81
Inquiry Question	inq_q	0.88	0.93	0.91
Instruciton	instr	0.70	0.09	0.15
Invite	inv	< 0.1	< 0.1	< 0.1
Inquiry Question	iq_q	0.83	0.35	0.50
Opinion	op	0.61	0.36	0.46
Other Question	otr_q	< 0.1	< 0.1	< 0.1
Prohibit	proh	0.53	0.19	0.28
Request	req	0.86	0.78	0.82
Request Question	req_q	0.90	0.72	0.80
Review	rev	0.45	0.06	0.10
Self_Intro	s_i	0.83	0.42	0.56
Suggestion	sug	0.84	0.13	0.23
Thank	thx	0.87	0.84	0.86
Urge	u	0.87	0.57	0.69
Warning	wn	0.93	0.53	0.68

TABLE VI. DIALOGUE ACT CLASSIFICATION SCORE FOR NAÏVE BAYES AND DIFFERENT KERNEL OF SVM CLASSIFIER

	Precision	Recall	F-score
SVC Linear kernel	0.768	0.781	0.758
SVC RBF kernel	**0.796**	**0.815**	**0.787**
SVC Poly kernel	0.771	0.753	0.727
MultinomialNB	0.728	0.729	0.696
BernoulliNB	0.73	0.739	0.733

V. CONCLUSION AND FUTURE WORK

In this work, we have explored the use of machine learning based MDAR, which can recognize the sequence of utterances in a conversation with multiclass DA classification on MmTravel corpus. Experimental results highlight that the SVM RBF kernel is better than the other approaches. Our perspective for the near future is to improve our MmTravel corpus from 60k to 100k, and also be aware of the dialogue act tag unbalancing data which will be overcome the zero F1 score for some act tags. Nowadays, every research has a scope for improvement with deep neural networks. As a future extension to this work, we are planning to analyze the performance of our MDAR with sequence learning neural networks.

APPENDIX A

1. (Because I have found fresh fruits here)
2. (Do you have a map?)
3. (I want to book an air ticket to Mandalay until next Friday)
4. (The best jewelry)
5. (Ok, I will take it)
6. (Could you please bring it to the laundry?)
7. (Put them in the car)
8. (Thanks for waiting)
9. (I apologize for being late)
10. (I can’t give you any money to buy clothes)
11. (It is better to arrive at the airport before 8 am)
12. (Good morning, how are you?)
13. (As far as I know, this is the best of all)
14. (Hope for the best)
15. (There is a shopping center next street)
16. (Now I have to sent another packet because of you)
17. (Do you want me to bring something?)
18. (Hot or cold)
19. (Don’t smoke, this is non-smoking area)
20. (Enjoy your flight!)
21. (Silent!)
22. (I’m Thuzar)
24. (Let me go, see you tomorrow)
25. (We always welcome you)
26. (This is the last time to you, do what I say next time)
27. (Why font is not working, it is inconvenient?)
28. (Really, congratulation)
29. (Why you didn’t buy?)

APPENDIX B

နောက်ထပ် ရော ဘာ ဖြစ်သေး လဲ (What else happened)
မင်း အချိန် ရှိလာ မှာ ပါ (You will have time)
ဒါလေး ယူသွား ပေးပါ လား (Please take this)
ခင်ဗျား စီးချင် တဲ့ လေကြောင်းလိုင်း ကို ခေါ် ပြီး ခင်ဗျား ဘာတွေ သယ်နိုင် လဲ ဆိုတာ မေးနိုင် ပါတယ် (Contact the airline you want to take and ask what you can bring)
အခုလို တွေ့ ရတာ ဝမ်းသာ ပါတယ် (Nice to meet you too)
ဒီလမ်း အတိုင်း တည့်တည့် လျှောက် ပါ ပြီးရင် ညာကွေ့ (Go straight and turn right)

ACKNOWLEDGMENT

We are grateful to Dr. Saw Myat Sandi Lwin, Myanmar Department, Hinthada University, for her assistance and suggestions on Myanmar dialogue act tag annotation. A special word of gratitude is due to the staff at Myanmar Department of the University of Yangon, who gave their time to discuss and advise us on various aspects of this research.

REFERENCES

[1] A. Lubna Shala1, Vasile Rus, and C. Arthur Graesser, "Automated Speech Act Classification in Arabic, " The University of Memphis.

[2] A. Grau, E. Sanchis, Mar´ıa Jos´e Castro and D. Vilar, "Dialogue act classification using a Bayesian approach," SPECOM, 9th Conference Speech and Computer, Russia.

[3] A. Stolcke and E. Shriberg (1998), "Dialog Act Modeling for Conversational Speech," AAAI Technical Report SS-98-01, 98–105.

[4] A. Stolcke, K. Ries, N. Coccaro, E. Shriberg, R. Bates, D. Jurafsky, P. Taylor, R. Martin, C. V. Ess-Dykema, M. Meteer, "Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech," Computational Linguistics, vol. 26, pp. 339–373, 2000.

[5] B. Gambäck, F. Olsson, O. Täckström, "Active Learning for Dialogue Act Classification," In Proceedings of the 12th Annual Conference of the International Speech Communication Association (INTERSPEECH), Florence, Italy, pp. 1329–1332, 27–31 August 2011.

[6] D. Jurafsky, E. Shriberg and D. Biasca, "Switchboard SWBD-DAMSL Shallow Discourse Function Annotation," (Coders Manual, Draft 13). Tech. Rep. 97-01, University of Colorado, Institute of Cognitive Science, 1997.

[7] D. Surendran and G.A. Levow, "Dialogue Act Tagging with Support Vector Machines and Hidden Markov Models," University of Chicago.

[8] E. Ribeiro, R. Ribeiro, D. M. de Matos, "The Influence of Context on Dialogue Act Recognition," arXiv 2015, arXiv:1506.00839.

[9] H. Bunt, "The DIT++ taxonomy for functional dialogue markup," AAMAS 2009 Workshop, Towards a Standard Markup Language for Embodied Dialogue Acts, pp. 13–24, 2009.

[10] H. Bunt, J. Alexandersson, J. Carletta, J. Choe, A. C. Fang, K. Hasida, K. Lee, V. Petukhova, A. Popescu-Belis, L. Romary, "Towards an ISO standard for dialogue act annotation," Seventh conference on International Language Resources and Evaluation (LREC'10), 2010.

[11] J. Carletta, "Unleashing the killer corpus: experiences in creating the multi-everything AMI meeting corpus," Language Resources and Evaluation, vol. 41, no.2, pp. 181–190, 2007.

[12] J. Carletta, A. Isard, S. Isard, J. Kowtko, A. Newlands, G. Doherty-Sneddon and A. Anderson, "The reliability of a dialogue structure coding scheme," Computational Linguistics, vol. 23, pp. 13–31, 1997.

[13] J. L. Austin, How to do things with words. Oxford university pres, 1975.

[14] J. Allen and M. Core, "Draft of Damsl: Dialog Act Markup in Several Layers," 1997.

[15] J. Allen, Natural Language Understanding (2nd Edition). [S.l.]: Addison Wesley, 1995.

[16] P. Boonkwan and T. Supnithi. 2013. Technical report for the network-based asean language translation public service project. Online Materials of Network-based ASEAN Languages Translation Public Service for Members, NECTEC.

[17] Y. Liu, "Using SVM and Error Correcting Codes for Multiclass Dialog Act Classification in Meeting Corpus," INTERSPEECH 2006, Pittsburgh, Pennsylvania, pp 1938-1941, September 17-21.

[18] ဒေါက်တာခင်အေး၊ အတ္ထဗေဒနိဒါန်း

[19] ဒေါက်တာအောင်မြင့်ဦး၊ လူမှုဘာသာဗေဒ မိတ်ဆက်

[20] ဒေါက်တာအောင်မြင့်ဦး၊ လူမှုဘာသာဗေဒ သဘောတရား

[21] မမြင့်မြင့်စန်း၊ မြန်မာဘာသာစကားတွင် ဘာသာစကားအသုံးနှင့် အနက်အဓိပ္ပာယ်ဆက်သွယ်မှု၊ ပါရဂူဘွဲ့အတွက် တင်သွင်းသော ကျမ်း၊ မြန်မာစာဌာန၊ ရန်ကုန်တက္ကသိုလ်

[22] မဥမ္မာစိုး၊ မြန်မာဘာသာစကားရှိ လမ်းညွှန်မှု အသုံးများ၊ ပါရဂူဘွဲ့အတွက် တင်သွင်းသောကျမ်း၊ မြန်မာစာဌာန၊ ရန်ကုန်တက္ကသိုလ်

[23] မောင်ခင်မင်(ဓနုဖြူ)၊ လက်တွေ့ အတ္ထဗေဒနိဒါန်း

[24] မောင်ခင်မင်(ဓနုဖြူ)၊ အတ္ထဗေဒနိဒါန်း