INTENT CLASSIFICATION OF USER COMMENTS IN MYANMAR LANGUAGE ON SOCIAL MEDIA SHOPPING PAGES

EI MYAT MYAT NOE

M.C.Sc **DECEMBER 2022**

INTENT CLASSIFICATION OF USER COMMENTS IN MYANMAR LANGUAGE ON SOCIAL MEDIA SHOPPING PAGES

BY

EI MYAT MYAT NOE

A Dissertation Submitted in Partial Fulfillment of the Requirement for the Degree of

Master of Computer Science

(M.C.Sc.)

University of Computer Studies, Yangon

DECEMBER 2022

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude and sincere thanks to all persons who have contributed directly or indirectly towards the completion of this thesis and helped me make this dissertation possible.

I would like to express my respectful thanks to Dr. Mie Mie Khin, Rector of the University of Computer Studies, Yangon, for her kind permission to submit this thesis.

My special thanks go to Dr. Yuzana, Pro-Rector of the University of Computer Studies, Pyay, for her administrative support, as throughout the development of thesis.

I would respectfully like to thank Dr. Aung Nway Oo, Professor, Faculty of Computer Science, University of Information Technology, for his valuable suggestion and encouragement of this thesis.

I would like to thank course coordinators, Dr. Si Si Mar Win, Professor, University of Computer Studies, Yangon and Dr. Tin Zar Thaw, Profesor, University of Computer Studies, Yangon for their superior suggestions and administrative supports during my academic study.

I am deeply thankful to my supervisor, Dr. Hsu Myat Mo, Lecturer, Faculty of Information Science, Natural Language Processing Lab, University of Computer Studies, Yangon, for her valuable guidance, encouragement, supervisor suggestions and supervision on the accomplishment of this thesis.

I would like to thank, Daw Hnin Yee Aung, Lecture, Department of English, University of Computer Studies, Yangon, for editing my thesis paper from language point of view.

I especially thank my parents, all of my colleagues, and friends for their encouragement and help during my thesis.

Finally, I would like to thank all of the staff, teachers from University of Computer Studies, Pyay for their support.

ABSTRACT

According to the social media usage statistics, there are about four billion total social media users across all platform. Moreover, people spend more time on social media than before and even perform daily activities on them. Because of the convenience of social media services, activities such as online shopping can be easily done on social media. Social media commerce has been one of the popular ecommerce trends in recent years. AI applications in customer services such as chatbots and personalization help business to understand their customers better and improve customer experience. Intent classification is one of the techniques used in these applications. This paper focuses on the classification of users' intentions based on the user comments posted in Myanmar Language on social media shopping pages. Convolutional Neural Network (CNN) is applied to classify the users' comments to one of the predefined intent categories. According to the experimental result, intent classification model with name normalization plus word segmentation preprocessing can give the F-score value of 86.6.

TABLE OF CONTENTS

Pages

LIST OF FIGURES

Pages

LIST OF TABLES

Pages

LIST OF EQUATIONS

Pages

CHAPTER 1

INTRODUCTION

Social media where people communicate, share thought and opinions about a particular topic, exchange information and express feeling, has made people's lives easier and has become an integral part of one's life. Social media enables communication for not only one's personal life but also for business life. People started adopting digital marketing instead of marketing offline and social media plays a vital role in promoting online business.

Most of the social media users usually post comments as their opinion about something, feedbacks of the services or products, and also as their inquiry about a particular information. Analyzing the users' generated content such as comments on social media posts has become an essential task to know the intentions behind those comments which can be very beneficial for the business.

Intent classification (sometimes called as intent recognition) is a technique used in Natural Language Processing (NLP) and the fundamental concept of Natural Language Understanding (NLU). Intent classification is a kind of text classification. Intent classification takes the written or spoken text as input for processing and uses machine learning (ML) and NLP techniques to analyze the intentions behind the input written or spoken texts and assigns or categorizes them into their predefined intent automatically.

AI applications in customer services such as chatbots and personalization help businesses to understand their customers better and improve customer experience and can also provide better services to their customers. Intent classification is one of the techniques used in these applications. It is useful to understand the intensions behind customer queries, automate process, and gain valuable insights. By understanding the users' intention, business can give a more accurate response to their users and customers.

Intent classifier can be modeled by applying traditional machine learning classification models such as k-Nearest Neighbor (KNN), Support Vector Machine (SVM), Naïve Bayes and so on. However, the diverse nature of contents on social media can make difficult for those traditional machine learning methods to classify intent classes. In recent years, deep learning models have proved that these models give impressive results on image processing and computer vision, speech recognition, NLP tasks such as machine translation, text classification, and many other tasks.

The filters/ kernel in CNNs can help identify relevant patterns in text data depending on kernel size. Since CNNs are translation invariant, they can detect the relevant patterns irrespective of their position in the sentence. Each filter/ kernel detects a specific feature, such as if the sentence contains some key phrases terms anywhere in the sentence. Most test classification tasks such as intent classification are determined by the presence or absence of some key phrases present anywhere in the sentence. Therefore, text classification tasks like intent classification task can be effectively modeled by CNNs which are good at extracting local and position-invariant features from data. Most comments on social media are written in free ordered and no regular standard grammar order. For this reason, CNN has been chosen to be applied for this intent classification task of users' comments in Myanmar text on social media shopping pages.

Myanmar language is free order language and the scripts are written from left to right without placing a regular space between words or phrases. Word segmentation step is one of the most important preprocessing tasks for Myanmar NLP research. Moreover, name words are also needed to be tokenized correctly because it can also affect the model accuracy. In this paper, before intent classification step, word segmentation and name normalization steps are performed as preprocessing tasks.

Without name normalization step but only with word segmentation processing step, the model accuracy is not as good as the model with word segmentation and name normalization preprocessing steps. Word level tokens are considered as input tokens to CNN model training for this intent classification task. The results show that CNN can give the promising result even though performed on the unbalanced and low amount of data.

In this dissertation, intent classification of users' comments in Myanmar language on social media shopping pages is carried out in deep learning approach by applying CNN. Moreover, the performance of two intent classification models is compared. As part of this intent classification, manually annotated corpus is created with the intention of to be able to use not only for this experiment but also for the future experiments.

1.1 Objectives of Thesis

The objectives of thesis are follows:

- To understand customer's intent and give a more accurate response to their customers
- To allow businesses to automate the interaction between interested buyers and business representatives
- To be able to apply in AI applications in customer services such as chatbots and personalization applications
- To be able to use in further Myanmar NLP and NLU research work

1.2 System Overview

This task is the classification of users' comments in Myanmar Language on Social Media platform. This task can be assumed as the early step of Natural Language Understanding (NLU) tasks. This kind of classification model can be applied in chatbots to provide appropriate answers to users. Humans can easily understand the meaning of different intents but chatbots system will require more complex techniques.

The experiments are carried out by applying the CNN, deep neural approach, and also investigated the effectiveness of the CNN. A manually intent annotated corpus is also developed to be used in machine learning approaches. From the experiments, the proposed intent classification model gives the promising result.

1.3 Organization of Thesis

This thesis is an implementation of intent classification of users' comments in Myanmar language on social media shopping pages. This dissertation consists of the following chapters:

Chapter 1 describes the introduction of the system, aims of thesis and introduction of intent classification.

Chapter 2 discusses various kinds of intent classification and practical use of intent classification and some literature review.

Chapter 3 presents the convolutional neural network and how they are implemented in this system. Moreover, description of Myanmar language is presented.

Chapter 4 is the process flow, detailed design and implementations of the system.

Chapter 5 consists of the advantages of the system, limitations and conclusion of the system.

CHAPTER 2

LITERATURE REVIEW

This chapter presents the various kinds of intent classification, practical use of intent classification and related works.

2.1 Intent Classification

Based on the advancement of technology, there are more businessmen who applied technology in their business with many useful ways. Providing high quality and modern services to customers is one of the things that businessmen and companies compete with each other. Intent classification, which is a part of Natural Language Understanding (NLU), is used to distinguished customers' attitudes, opinions and intentions from such services.

The purpose of using Natural Language Understanding (NLU), a subset of Natural Language Processing (NLP) is to check the content of word, the grammar of those words and improve machine understanding. By using chatbots and AI technologies, business owners and company owners work to fulfill the customers' intent.

By using Natural Language Processing (NLP), text classification or text categorization, the process of categorizing text into organized groups, is one of the important techniques for business. Text classifiers can automatically analyze text and then assign a set of pre-defined tags or categories based on its content. Using text classifiers, companies can automatically structure all manner of relevant text, from emails, legal documents, social media, chatbots, surveys, and more in a fast and cost-effective way. This allows companies to save time analyzing text data, automate business processes, and make data-driven business decisions.

Intent classification (focused on future action) is a form of text classification [8]. However, it is different from the well-studied problems of topic classification [9] (focused on matter) as well as subjective text classification such as sentiment or emotion classification (focused on the current state of affairs) [10].

Intent classification is a kind of text categorization technique. Intent classification is a machine-learning approach that groups text into pre-defined intent categories. It is an integral tool in Natural Language Processing (NLP) used for varied tasks like spam and non-spam email classification, sentiment analysis of movie reviews, detection of hate speech in social media posts, etc.

There are a lot of machine learning algorithms available for text classification like Naïve Bayes, Support Vector Machines, Logistic Regression, etc. All these machine learning algorithms can be used for Intent classification. Moreover, intent classification problems can be solved by applying deep learning approaches, and fine-tuning approach.

Deep learning based Convolutional Neural Network (CNN) architecture is used for this intent classification of users' comment posted in Myanmar language on social media pages.

2.2 How does Intent Classification work?

In Machine Learning and Natural Language Processing, intent classification is used to describe words and expressions. For example, a machine learning model can learn that words such as buy or acquire are often associated with the intent to purchase.

However, intent classifiers need to be trained with text examples first, otherwise known as training data. When analyzing customer emails, the predefined tags such as, *Interested, Need Information, Unsubscribe, Wrong Person, Email Bounce, Autoreply* might be chosen.

The input can be in the form of text or speech. Speech input must be converted to text form with speech-to-text technology. Intents that are appropriate for the use case should be determined. Then, intent classifiers are trained and text is labeled with intents. For example, the sentence "I want to chat with David" indicates a request and can be tagged as such.

With the defined tags, the intent classifier can be trained for the relevant text examples for each tag. For example sentence is "I tried to make a purchase through the site but I don’t know where to start, could you help me out?” This email can be tagged as *Interested*.

2.3 Why is Intent Classification useful?

Intent classification allows businesses to be more customer-centric, especially in areas such as customer support and sales. From responding to leads faster , to dealing with large amounts of queries and offering a personalized service, intent classification can be a key tool.

Intent classification is used in conversational AI applications to provide personalized conversation experiences to users. It helps to increase sales and improve overall customer experience. It allows businesses to understand customers’ intent and give a more accurate response to their customers.

It can allow businesses to automate the interaction between interested buyers and business representatives. It classifies customers’ needs and their special attention then analyzes what customers intend to achieve, and directs them to the relevant representative by categorizing customers’ intents.

Automatically linking words or sentences with specific intent is achieved by intent classification, which combines machine learning and natural language processing. A machine learning model, for instance, can discover that phrases like “purchase” or “acquire” are frequently linked to the intention to buy.

Intent classification works to enhance, streamline and systemize the customer service process. Without intent classification software, customer service agents have to sort through customer messages, answer phone calls manually and monitor-based instant messaging platforms to understand the intent behind each query.

With intent classification software, the trained algorithm analyzes, tags and assigns messages based on intent. The software will escalate messages to customer service agents or assign tasks to chatbots when possible. As a result, both customer service representatives and customers can enjoy benefits, such as:

1. Improved Customer Experience
2. Consistent Care
3. Increased Conversions

2.4 Why does Intent Classification Matter to Customer Service?

Intent classification is important because customers' intent are correctly identifying is lead to a much quicker and more friction-free experience for the customer (or involving agents). Agent time is not wasted on poorly directed customer queries, and more importantly customer time is not wasted.

Efficient intent classification does not just figured out what customer needs. Automatically and quickly detecting intents to purchase is crucial for sales and customer support because it allows companies to take immediate action and transform leads into paying customers. The faster teams can detect purchase intents and respond, the better their chances of closing a deal.

Prospective clients appreciate fast-paced responses, with some expecting a response in less than 6 hours. When the clients receive a Facebook message asking for product availability, with an intent classifier, the clients can quickly identify this interaction as an interested client and contact them immediately to boost the possibility of turning this engagement into a sale.

Even when companies are inundated with data, intent classifiers are able to pinpoint customers who have expressed interest and direct these specific queries to the sales teams. Machines work faster than humans, non-stop, and they do not grow tired so even as workloads increase, they will never miss a potential sale.

Machines are consistent in the way they operate, preprocessing data using the same parameters and criteria for every single instance. The consistency in criteria ensures that all the customer intents are analyzed under the same circumstances, applying the same standards, protocols, algorithms, etc. This greatly reduces errors and improves data accuracy.

As a marketing campaign deployed and started receiving customer interactions, intent classifiers can be used to identify potential customers that show a "high intent" to purchase and contact them immediately. Thus, the customers' conversion rates skyrocket.

When clear intents identified automatically in sales and marketing campaign, the client could easily create reports based on factual data about conversion rates, interested buyers, upsell opportunities, and much more.

2.5 Related Works

The authors of [22] considered the task of annotating travel-related reviews with travel intents that best represent the reviewer's reasons for visiting the place of interest. They classified the reviews into eight travel intent, business, eating out, education, health, holidays, and religion, shopping and socializing. They applied Naïve Bayes as a baseline method and compared with other classification models such as DNN, SVM and RF.

The authors of [4] addressed the problem of multiclass classification of intent with a use-case of social data generated during crisis events. The crisis data was classified into three different intents, seeking, offering, and none (neither seeking nor offering). The hybrid approach that combines knowledge-guided patterns with syntactic features based on bag of tokens.

According to the authors of [1], text classification is a representative research topic in the field of natural language processing that categorizes unstructured text data into meaningful categorical classes. They classified the data using the Internet Movie Database (IMDB) movie review data to evaluate the performance of the proposed model, and the test results showed that the proposed hybrid attention Bi-LSTM and CNN model produces more accurate classification results. The proposed model achieved higher accuracy which increased as the size of training data and the number of training epochs increase.

The authors of [8] proposed a simple and novel multi-point semantic representation framework with relatively low annotation cost to leverage the fine-grained factor information, decomposing queries into four factors, i.e., topic, predicate, object/condition, query type. The model significantly out-perform several state-of-the-art approaches with an improvement of 1.35%-2.47% in terms of accuracy.

Intent classification and slot filling are two essential tasks for natural language understanding [17]. They often suffer from small-scale human-labeled training data, resulting in poor generalization capability, especially for rare words. So, this paper used a joint intent classification and slot filling model based on BERT. Experimental results demonstrate that their proposed model achieves significant improvement on intent classification accuracy, slot filling F1, and sentence-level semantic frame accuracy on several public benchmark datasets, compared to the attention-based recurrent neural network models and slot-gated models.

This paper considered the diversity in text representation; traditional machine learning has been unable to accurately understand the deep meaning of user texts. The proposed system used a BERT pre-trained model in deep learning based on Chinese text knots, and then added a linear classification to its. This paper performed domain intent classification experiments in the Chinese texts THUCNews Dataset. Compared with Recurrent Neural Network and Convolutional Neural Network methods, this method can improve performance by 3% points. Experimental results of BERT pre-trained model can provide better accuracy and recall of Chinese news text domain intent classification [16].

The authors of [11] proposed a simple and general method to regularize the fine-tuning of Transformer-based encoders for text classification tasks. During fine-tuning, the model generate adversarial examples by perturbing the word embedding matrix of the model and perform contrastive learning on clean and adversarial examples in order to teach the model to learn noise-invariant representations. For the challenging low-resource scenario, the system using half of the training data (per intent) in each of the three intent classification datasets, and achieve similar performance compared to the baseline trained with full training data.

The authors of [14] solved the problem of a Chinese medical intent dataset (CMID) using the questions from medical QA websites. The intent annotation is four types and 36 subtypes of users' intents. CMID also provides two types of additional information, including word segmentation and name entity. Among Fast Text, TextCNN, TextRNN and TextGCN, Fast Text and TextCNN models have achieved the best results in four types and 36 subtypes intent classification, respectively.

The authors of [13] critically investigated the available datasets for citation intent and proposed an automated citation intent technique to label the citation context with citation intent. Data are annotated ten million citation contexts with citation intent from Citation Context Dataset (C2D) with the help of proposed method. Global vectors (Glove), Infersent, and Bidirectional Encoder Representations from Transformers (BERT) word embedding methods are applied and compared their

Precision, Recall, and F1 measures. It was found that BERT embedding performs significantly better, having an 89% Precision score.

The authors of [9] proposed a dialog system that captures proper intent and activated slots for Korean in-vehicle services in a multi-tasking manner. The model was implemented with pre-trained language model and it includes an intent classifier, slot classifier, slot value predictor and value-refiner. The experiments are conducted on the Korean in-vehicle services dataset and show 90.74% of joint global accuracy. Also, the efficacy of each component of the model was analyzed and inspected the prediction results with qualitative analysis.

The authors of [10] compared the performance of the Japanese BERT model, one of the latest natural language processing technologies, with Word2vec, one of the conventional methods. Data are used from the Livedoor news corpus for the experiments. FAQ chatbot is built and compared the rate of correct answers to questions about news articles asked by the users between BERT and Word2vec. BERT showed superior performance compare to Word2vec.

This paper focuses on the intent classification of users' generated comments on social media posted in Myanmar text. Comments are annotated into five intent categories*:* application, auto-subscription, bill, customer service and internet. The proposed CNN model achieved over 0.94 F1-Score [21].

The authors of [2] attempted to analyze the Korean sentence classification system. Sentence classification is the task of classifying an input sentence based on predefined categories. This paper purposed a novel approach of Integrated Eojeol Enbedding to reduce the effect of poorly analyzed morphemes on sentence classification. This paper also proposed two noise insertion methods that further improve classification performance. The evaluation results indicate that by applying the proposed methods on the existing sentences classifiers, the sentence classification accuracy on erroneous sentences is increased by 8% to 15%.

The authors of [7] focused on optimistic machine learning and feature set selection to classify collected tweets. The system build the classifier model using Naïve Bayes and Naïve Bayes Multinomial, Support Vector Machine and Decision Tree Algorithms, all of which show good performance. Experiments show that precision and F-measure performance are best when using a Naïve Bayes Multinomial classifier model with a test feature set defined by extracting Substantive, Predicate, Modifier, and Interjection parts of speech.

The authors of [20] explored the effectiveness of two separate families of Deep Learning networks for those tasks: Bidirectional Long Short -Term networks and Transformer-based networks. The models were trained and tested on the ATIS benchmark dataset for both English and Greek languages. The proposed of this paper is to present a comparative study of the two groups of networks for both languages and showcase the result of the experiments. The model, being the current state-of-the-art, yielded impressive results and achieved high performance.

Data augmentation methods ameliorate this issue, but the quality of the generated data varies significantly across techniques. The authors studied the process of systematically producing pseudo-labeled data given a small seed set using a wide variety of data augmentation techniques, including mixing methods together. Finally, the mixing methods are likely to produce strong results [3].

Chapter Summary

In this chapter, the introduction about intent classification, how it works and why it is important and where it can be applied are described. Moreover, approaches to intent classification are also described and some intent classification research that have been done are discussed.

CHAPTER 3

BACKGROUND THEORY

The detail explanation of convolutional neural network, how these networks are implemented in this system and description of Myanmar Language is presented in this chapter.

3.1 Convolutional Neural Network

In deep learning, a convolutional neural network (CNN or ConvNet) is a class of artificial neural network (ANN), most commonly applied to analyze visual imagery. CNNs are also known as Shift Invariant or Space Invariant Artificial Neural Networks (SIANN), based on the shared-weight architecture of the convolution kernels or filters that slide along input features and provide translation-equivariant responses known as feature maps.

Counter-intuitively, most convolutional neural networks are not invariant to translation, due to the down-sampling operation they apply to the input. They have applications in image and video recognition, recommender systems, image classification, image segmentation, medical image analysis, natural language processing, brain-computer interfaces, and financial time series.

CNNs are regularized versions of multilayer perceptrons. Multilayer perceptrons usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The “full connectivity” of these networks make them prone to over-fitting data.

Typical ways of regularization, or preventing over-fitting, include: penalizing parameters during training (such as weight decay) or trimming connectivity (skipped connections, dropout, etc.). CNNs take advantage of the hierarchical pattern in data and assemble patterns of increasing complexity using smaller and simpler patterns embossed in their filters. Therefore, on a scale of connectivity and complexity, CNNs are on the lower extreme.

Convolutional networks were inspired by biological processes in that the connectivity pattern between neurons resembles the organization of the animal visual cortex. Individual cortical neurons respond to stimuli only in a restricted region of the visual field known as the receptive field. The receptive fields of different neurons partially overlap such that they cover the entire visual field.

CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns to optimize the filters (or kernels) through automated learning, whereas in traditional algorithms these filters are hand-engineered. This independence from prior knowledge and human intervention in feature extraction is a major advantage.

3.2 CNN Architecture

A convolutional neural network consists of an input layer, hidden layers and an output later. In any feed-forward neural network, any middle layers are called hidden because their inputs and outputs are masked by the activation function and final convolution. In a convolution neural network, the hidden layers include layers that perform convolutions.

The convolution neural network typically includes a layer that performs a dot product of the convolution kernel with the layer's input matrix. This product is usually the Frobenius inner product, and its activation function is commonly ReLU. As the convolution kernels slides along the input matrix for the layer, the convolution operation generates a feature map, which in turn contributes to the input of the next layer. This is followed by other layers such as pooling layers, fully connected layers, and normalization layers.

To explain convolutional neural networks in simple terms - just as parents train their children, computers are also trained by showing a million images of the same object so that their ability to recognize that object increases with each sample.

3.3 Structure of CNNs

CNNs are structured differently as compared to a regular neural network. In a regular neural network, each layer consists of a set of neurons. Each layer is connected to all neurons in the previous layer. The way convolutional neural networks work is that they have 3-dimensional layers in a width, height, and depth manner. All neurons in a particular layer are not connected to the neurons in the previous layer. Instead, a layer is only connected to a small portion of neurons in the previous layer.

3.3.1 Convolutional Layers

The top layer is perceived as the mathematical layer. It is essentially the convolutional layer and deals with understanding the number pattern it sees. Let's assume the first position in this layer starts applying a filter around the top left corner of the image. The filter is also referred to as a neuron or a kernel. It reads that part of the image and forms a conclusion of an array of numbers, multiplies the array, and deduces a single number out of this process.

- This single number represents the top left corner that the convolutional layer has just read of the image. The part of the image that the filter scans over is the receptive field. The filter then moves right by 1 unit and starts the same process again. In this fashion, the convolutional layer reads the entire image and assigns a single number to each unit. This data gets stored in a 3D array. In essentiality, this entire process functions like the human brain. The receptive field in the world of CNNs is the visual field in the world of human biology. The filter acts as the visual cortex containing small regions of cells targeting the reading of specific areas of the visual field.
- Suppose that we have some N*N square neuron layer which is followed by convolutional layer. If the input sentences have m*n the filter ω, the convolutional layer output size is (N-m+1) * (N-m+1). To calculate the pre-nonlinearity input unit x_{ij}^{l} in this layer, these units are sum up weighted by the filter components from the previous cells. In forward propagation,

$$x_{ij}^{l} = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} \omega_{ab} \quad y_{(i+1)(j+b)}^{l-1} \qquad \text{Equation (3.1)}$$

Then, nonlinearity functions are used to convolutional layer:

$$y_{ij}^{l} = \sigma(x_{ij}^{l}). \qquad \text{Equation (3.2)}$$

In backward propagation, for the convolutional layer, the weights are computed and propagated errors back to the previous layer. The chain rule equation is used:

$$\frac{\partial E}{\partial y_{ij}^{l-1}} = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} \frac{\partial E}{\partial x_{(i-a)(j-b)}^{l}} \frac{\partial x_{(i-a)(j-b)}^{l}}{\partial y_{ij}^{l-1}} = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} \frac{\partial E}{\partial x_{(i-a)(j-b)}^{l}} \omega_{ab}$$

Equation (3.3)

3.3.2 Pooling Layers

Convolutional networks may include local and/or global pooling layers along with traditional convolutional layers. Pooling layers reduce the dimensions of data by combining the outputs of neuron in the next layer. Local pooling combines small clusters, tilling sizes such as 2*2 are commonly used. Global pooling acts on all the neurons of the feature map. There are two common types of pooling in popular use: max and average. Max pooling uses the maximum value of each local cluster of neurons in the feature map, while average pooling takes the average value.

3.3.3 Fully Connected Layers

Fully connected layers connect every neuron in one layer to every neuron in another layer. It is the same as a traditional multilayer perceptron neural network (MLP). The flattened matrix goes through a fully connected layer to classify the images. This layer is the completion layer in a convolutional neural network. It takes the final output of the layer before it (be it a ReLU or a convolutional layer) and provides an N-dimensional vector output. ‘N' here signifies the number of classes the program chooses from. For example, if the program is looking at pictures of horses, it will look at high-level features such as 4 legs, hooves, the tail, or muzzle. This fully connected layer will look at the high-level features and connect that with the image thus giving the output of a classification of a horse.

3.4 CNNs Integrated with Deep Learning

Companies may find it difficult to integrate convolutional neural networks and neural networks into production-ready applications. Multiple factors need to be taken into consideration to make this happen, such as -

- What convolutional architecture should be used
- What kind of data and data sets should be used
- Which data model or deep learning model should be used to accommodate the data

3.4.1 Real World Applications of CNNs

Companies are usually on the lookout for a convolutional neural networks guide, which is especially focused on the applications of CNNs to enrich the lives of people.

Simple applications of CNNs which the user can see in everyday life are obvious choices, like facial recognition software, image classification, speech recognition programs, etc. These are terms that the laymen, which are familiar with, and comprise a major part of our everyday life, especially with image-savvy social media networks like Instagram. Some of the key applications of CNN are listed here –

- Decoding Facial Recognition
- Facial Emotion Recognition
- Object detection
- Analyzing Documents
- Auto Translation
- Next word prediction in Sentence
- Handwritten Character Recognition
- X-ray image Analysis
- Cancer Detection
- Visual question answering
- Image Caption
- Biometric Authentication
- 3D Medical Image Segmentation

- Understanding Climate
- Understanding Gray Areas
- Advertising
- Others Interesting Fields

3.5 Myanmar Language

Myanmar language (also known as Burma) it is an official language and the native language of the Myanmar, the country's principal ethnic group. Myanmar language is spoken all over the world in millions people. Myanmar language is written in left to write without spacing between words or syllables. Myanmar characters have three groups: 33 consonants (called 'Byee'), four basic medials (called 'Byee Twe') and vowels (called 'Thara').

33 Consonants	က ခ ဂ ဃ င စ ဆ ဇ ဈ ည ဋ ဌ ဍ ဎ ဏ တ ထ ဒ ဓ န ပ ဖ ဗ ဘ မ ယ ရ လ ဝ သ ဟ ဠ အ
4 Medials	ျ ြ ွ ှ
12 Vowels	ာ ါ ိ ီ ု ူ ေ ဲ ံ ့ း ်
Myanmar Digits	၁ ၂ ၃ ၄ ၅ ၆ ၇ ၈ ၉ ၁၀

Table 3.1 Example of Myanmar Alphabet

Word segmentation is a vital step for Myanmar NLP tasks. Because of the writing style of Myanmar Language in which there is no standard break between word boundary, it has become one of the challenges in Myanmar NLP. Therefore, word segmentation is one of the problems that is needed to be addressed in data processing step.

Furthermore, to recognize named entities correctly in sentences are also vital in Myanmar NLP. These all can affect the performance of the model results.

Chapter Summary

In this chapter, the explanations about CNN as well as how each layer of it works are described. Furthermore, a little explanation of Myanmar language and some challenges have to be faced before Myanmar text processing is discussed.

CHAPTER 4

DESIGN AND IMPLEMENTATION

This chapter describes the design and implementation of intent classification model using Convolutional Neural Network.

4.1 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is a type of artificial neural network [5]. Although CNNs are mainly used in image processing and computer vision problems, lately they have been applied in NLP tasks with interesting outcomes.

In this proposed CNN model, there are four layers: embedding layer, convolutional layer, pooling layer, and fully connected output layer. The incoming input sentences are needed to convert into dense vectors before further processing. And then the result vector is described as vector representation. From the convolutional layer, various features form the input can be extracted. This layer is the first layer of Convolutional Neural Network (CNN) model.

By sliding the filter and the parts of the input with respect to the size of the filter, the output is termed as the feature map which gives the information about the sentence. Later, this feature map is fed to the other layer to learn the several features of input. The next layer is the pooling layer that comes after the convolutional layer.

The pooling layer reduces the size of the feature map which means it only retains the important features of the input. Pooling can be performed either average pooling or max pooling. In max pooling, the largest element is taken from the feature map. Average pooling calculates the average of the element.

The last layer is the fully connected layer. In the fully connected layer, an activation function is used to calculate the probability of each input to every output vector. Figure 4.1 shows the CNN model architecture for intent classification of is presented.

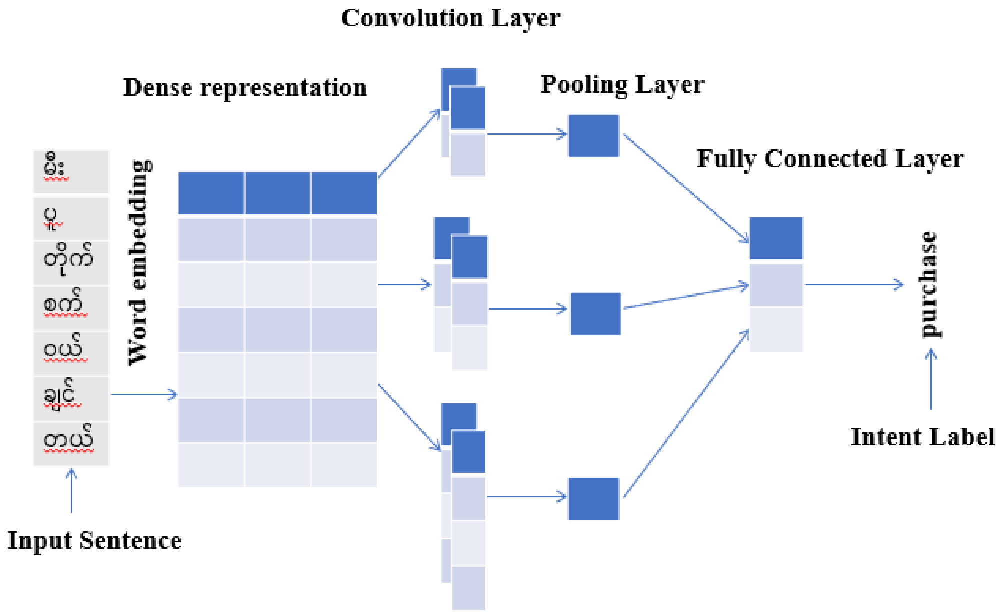

Figure 4.1. CNN Architecture for this Intent Classification Model

4.2 Intent Classification Model

In this section, the overall work flow of the proposed CNN intent classification modeling will be described. Figure 4.2 shows the work flow of the proposed intent classification modeling. Firstly, Myanmar language is a low resource language which means there no efficient linguistic resources such as data that are needed to process with machine learning algorithms.

As a very step for this intent classification model, a corpus was manually constructed. In the corpus building stage, data are firstly collected from Shopping Mall pages where the users' comments in Myanmar Language. These comments are retrieved from like Terminal M[1], Ocean[2], Junction Square[3], Sein Gay Har[4], Gamone Pwint[5], Time City[6] and many other shopping center and online shopping pages like

[1] https://www.facebook.com/terminalm.ygn
[2] https://www.facebook.com/OceanSupercenterMyanmar
[3] https://www.facebook.com/junctionsquareshoppingcentre
[4] https://www.facebook.com/SeinGayHar
[5] https://www.facebook.com/gamonepwintshoppingmalls
[6] https://www.facebook.com/TimesCityYangon

cosmetic online pages and fashion online pages, online shopping pages that sell consumer goods and other online shopping pages.

Users' comments are preparing in font converting and spelling errors. Data preprocessing steps (word segmentation and name normalization) are performed on this manually prepared corpus before training process. After that intent classification model training is conducted with CNN neural architecture.

In decoding stage, the obtained CNN based intent classification model is used to classify the user input comments into their corresponding intent categories.

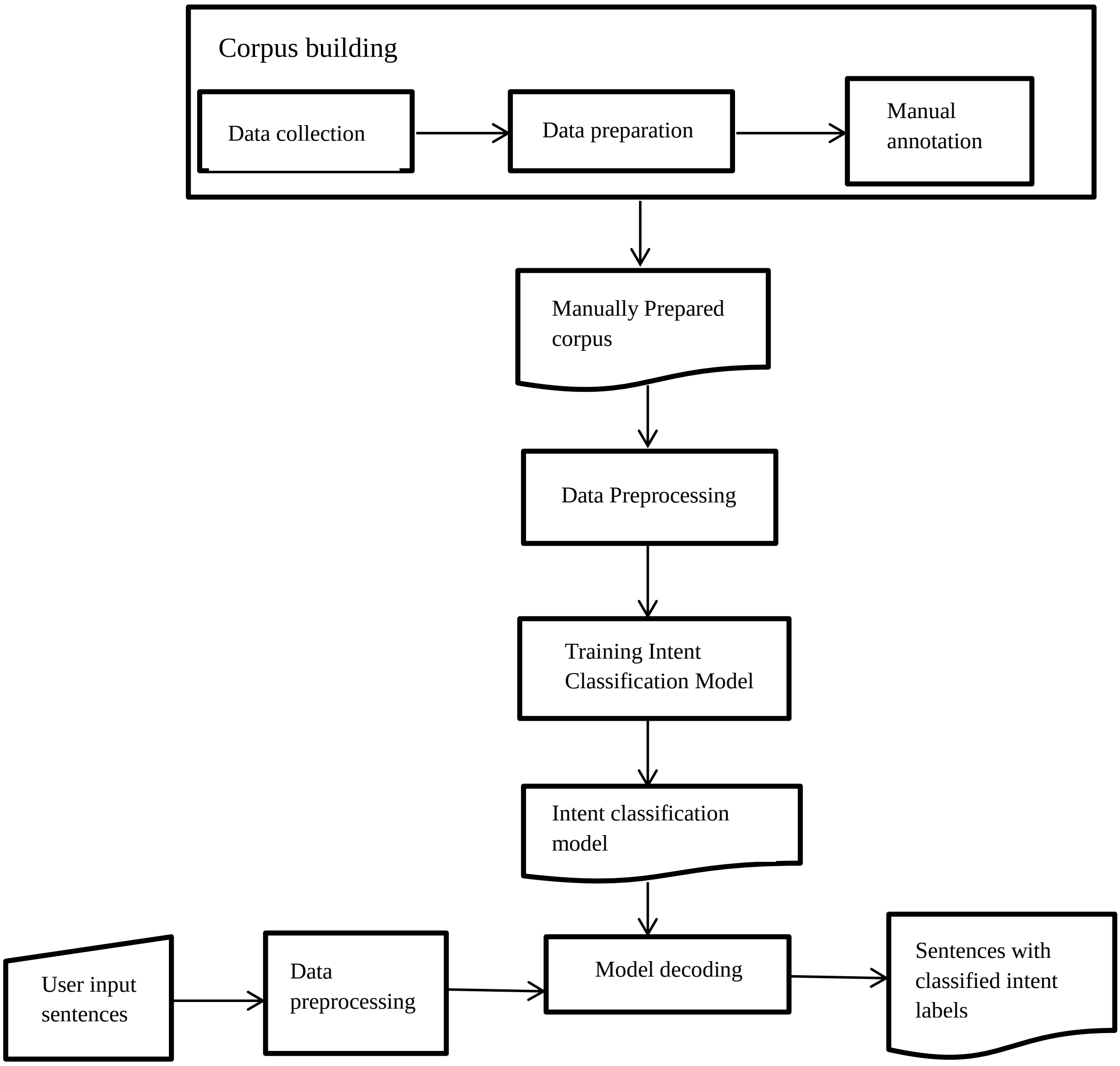

Figure 4.2. Work Flow of the System

4.2.1 Corpus Building

There are many vendors and shopping malls in Myanmar and most of them have their own official social media pages (Facebook pages). Users' comments written in Myanmar language from these pages are collected as raw data to build the corpus.

It is important to prepare the data before any other tasks. Comments on social media are written in different fonts. Firstly, all the collected data are converted from non-Unicode fonts such as Zawgyi font to Unicode font in order to get the unique Unicode encoding, and are corrected spelling errors and mistyped errors in text. Moreover, noise data are cleaned; such as by eliminating emojis from the sentences, and all the foreign words are transliterated.

Font encoding is one of the problems to be addressed before any other tasks in Myanmar NLP research. Most of the users are familiar non-Unicode font encoding and these fonts are mostly occurred in social media content. To be unique and to be able to use in text computation, all the data are needed to convert to Standard Unicode encoding. In this work, Rabbit Converter[7] is used to convert Zawgyi to Unicode. As an example, the following sentence with Zawgyi encoding: "အဲဆိုင္ အသစ္ဖြင့္တာတဲ့" is converted as "အဲဆိုင် အသစ်ဖွင့်တာတဲ့".

Moreover, social media content data are noisy. Thus, it is necessary to check spelling errors and mistyped errors in the data. For example, the character "ဝ" (WA) is usually wrongly typed as "၀" (Zero). Another most commonly occurred errors are the characters such as "ွ့, "း", and "် are missing or duplicated in sentences. Besides this noisy nature, social media comments are usually posted with emojis. It is also needed to eliminate form text for text processing. For example, after eliminating the emojis from the following sentences "ဒီနေ့ လူတွေကိုစည်ကားလို့။😊😊 ", the cleaned sentence: "ဒီနေ့ လူတွေကိုစည်ကားလို့။" is ready for text processing. And the next step

[7] https://www.rabbit-converter.org/Rabbit/

is transliterated process that performed English names are transform into Myanmar language.

After being performed these data preparation tasks, all the prepared data are manually annotated with the predefined seven types of intent categories. In this work, seven different types of intent that are related to the shopping activities are defined to classify the users' comments.

The intent "purchase" is used to annotate the comments that are written with the intention of buying a product or something from the vendors. The intent "payment" is for comments that are related to the action or process of paying for the products. The intent "delivery" is used to annotate the users' comments that are related to the delivery services. The intent "job" is used when the comments appear like want to apply for job or inquiry for job position. The intent "opinion" is applied when the users' comments seek like expressing their opinion about the products or services.

When users' comments are questions asking, they want to know something about the services or products, these texts are annotated with the intent "general_info". The next intent "contact_info" is considered for the annotation of users' comments that are posted with the intention of inquiring the contact information such as address or phone no of the vendors or services and where they can buy the products. Table 4.1 shows the examples of defined intent categories and their usage in the annotation process.

There are over 1K sentences in this manually annotated intent corpus. Table 4.2 shows the data statistic and distribution of each intent category in this manually annotated corpus. It contains the intent category "general_info" the most; over 27% over the whole dataset; and the intent category "job" is the least; this type of intent is about 5% in the dataset.

Table 4.1. Description of the defined intent classes and sample data

No	Defined Intents	Example sentences
1	Purchase	မီးပူတိုက်စက်ရှိလားရှင့်။
2	Payment	ဆိုင်မှာလာဝယ်ရင် ကေပေး နဲ့ရှင်းလို့ရလားရှင့်။
3	Job	ဝန်ထမ်းခေါ်လားဗျ။
4	Contact_info	ဘိုကလေးဈေး ဂမုန်းပွင့် ဖုန်းနံပါတ်သိချင်ပါတယ်
5	Delivery	နယ်ကမို့ပါ။ လူကိုယ်တိုင် မလာဝယ်နိုင်တဲ့အတွက် စာတိုက်ကို ပို့ပေးနိုင်ပါသလားရှင့်။
6	General_info	ကားပါကင်က တစ်နာရီကို တစ်ထောင်လားရှင့်။
7	opinion	ကောင်းလိုက်တဲ့ ဝန်ဆောင်မှုပါ။

Table 4.2. Data statistic and data distribution of dataset

Data	Size	%
Total data	10335	
Purchase	2165	20.948%
Payment	586	5.670%
Job	158	1.529%
Contact_info	1196	11.572%
Delivery	1837	17.775%
General_info	2858	27.654%
opinion	1535	14.852%

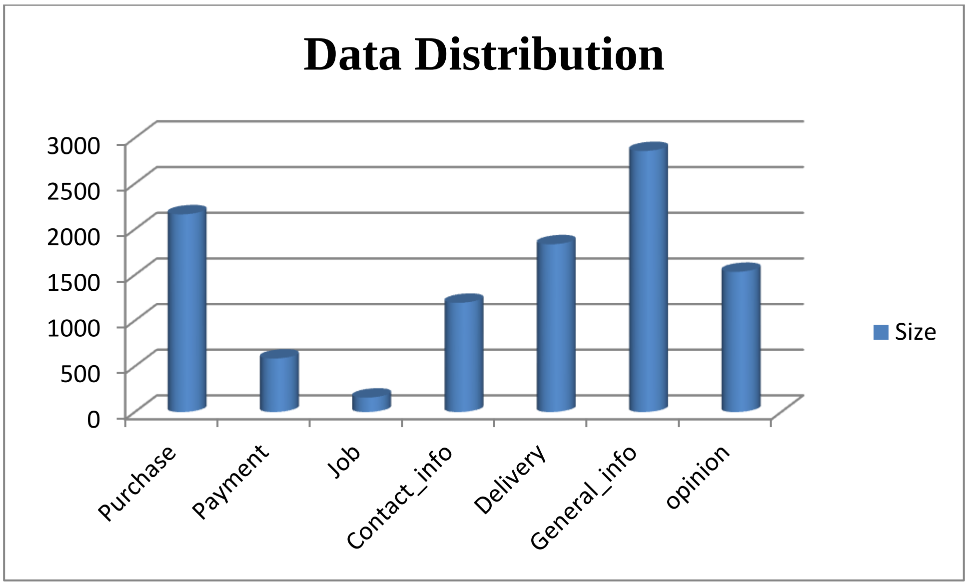

Figure 4.3. Data statistic and data distribution of dataset as described in graph

4.2.2 Preprocessing

Myanmar scripts are written in free order and are also written continuously in sequence with or without breaking between words which means that there is no standard placement of white space between words or phrases in sentences. Therefore, useful information about words that are very important to text processing with machine cannot be got. Word segmentation is a very important preprocess for Myanmar NLP tasks. In this work, word level is considered as basic input tokens to the CNN input layer. For this word segmentation process, users' comments are segmented into words by using word-segmentor[8] for Myanmar language from UCSY-NLP research lab. Figure 4.4 shows the raw data before (word segmentation) and figure 4.5 shows word segmented data by using the word-segmentor.

ဘိုကုန်းပို့ခဘယ်လောက်လဲရှင့်

Figure 4.4. Raw Data before (word segmentation)

[8] http://www.nlpresearch-ucsy.edu.mm/wordsegmentation.html

ဘို ကုန်း ပို့ ခ ဘယ်လောက် လဲ ရှင့် ။

Figure 4.5. Word Segmentation Data by using Word-Segmentor

Moreover, many name words appear in data and it is necessary to correctly recognize the name word in text. It can also affect the classification accuracy. For this reason, named normalization process is carried out to correctly segment name words and recognize names in text. For this name normalization process, named entity recognizer for Myanmar language from UCSY-NLP research lab [6] was firstly applied to recognize names in comments. Figure 4.6 shows the named recognized data by using name entity recognition (NER). After that, named entities are normalized with respective name words.

ဘိုကုန်း ပို့ခ ဘယ်လောက် လဲ ရှင့်

Figure 4.6. Named Recognized Data by using Name Entity Recognition (NER)

During intent classification modeling, two different types of experiments are carried out: the first one is carried out with data that are only preprocessed with word segmentation process; and another is carried out with data that are obtained from both word segmentation and named normalization processes. The experimental results show that as data preprocessing step, named normalization is as important as word segmentation for the text processing.

4.2.3 Training Set Up

For the CNN neural training for this intent classifier modeling, TensorFlow[9], Keras[10] deep learning python framework is used. All the experiments are conducted on Google Colab service provided by Google. Data is partitioned into 80% for

[9] https://www.tensorflow.org/

[10] https://keras.io/

training and 20% for testing. Table 4.3 shows the data partition for training and testing.

Training data are passed to the embedding layer of CNN for dense vector representation of each input word. Embedding dimension size is set as 60. Input maximum length is 10.

One-dimensional filters of size 3 convolve over one-dimensional input. 64 filters are applied and for each filter, values of input word vectors are pairwise multiplied with associate weights in the filter and then results are being summed up and passed into ReLU activation function to produce output feature maps.

As for the pooling layer, pooling is performed with one-dimensional max-pooling to capture the important features. Outputs from pooling layer are flattened and passed into fully connected output layer to predict the probability of each intent class. In this layer, softmax is applied as activation function.

Table 4.3. Data partition for training and testing

Data	**Size**	**%**
Total Data	10335	
Training data	8267	80%
Testing data	2067	20%

4.3 Experimental Result

Performance of the trained models is measured with the F1-score. Firstly, intent classification model is trained and tested with data that are preprocessed by only word segmentation. This model gives the F1-socre of 0.70%. The proposed model, which is trained with the named normalized and word segmented data, outperforms; and F1-score value of 0.86% is obtained. Comparison of F1-score for each model is shown in Table 4.4.

Although CNN is effective when the data is well prepared and trained on very large amount of data, from this experiment, it can be seen that it still works properly while training on the low amount of data if the data is well prepared.

Therefore, more data is needed to get the full effectiveness of CNN. However, the experimental result is promising.

Table 4.4. Comparison of F1-score for each model

Model	**F1-score**
With word segmentation	0.70
With word segmentation and name normalization	0.86

Chapter Summary

In this chapter, detail explanation of how to implement the intent classifier to classify users' comments is described. Starting from the data creation, to model training, each process is presented step by step. Finally, evaluation results are compared and presented in accordance with the experimental result.

CHAPTER 5

CONCLUSION

Intent classification allows businesses to be more customer-centric, especially in areas such as customer support and sales and it allows business to understand customers' intent and give a more accurate respond to their customer. From responding to lead faster, to deal with large amounts of queries and offering a personalized service, intent classification can be a key tool. The proposed system will help in classifying customers' needs and their special attention as well as in analyzing what customers intend to achieve. In this work, a deep-learning-based convolutional neural network architecture is applied to perform intent classification of users' comments written in Myanmar language.

Moreover, a manually intent annotated corpus with the size of over 1K is also created which can be a very useful linguistic resource for future Myanmar NLP research.

Two types of intent classification models are presented and compared their performance. The first model is obtained from training on word segmented data. Another model which is the proposed model of this dissertation is trained with the data obtained from two preprocessing processes (word segmentation and named normalization). In data many name words appear, thus name normalization is also applied as one of the preprocessing steps before training and it is more efficient than the model trained with only word segmentation process. According to the experiments, the F1 score of the model (with word segmentation and NER) is about 15% greater than that of the model with only word segmentation.

5.1 Limitations

Although the F1 score of the proposed intent classification model is not as good as other intent classification model, it is still promising. According to the testing results, ambiguity of intent categories is the most seen error. Because of the nature of imbalanced data, most of the users' input test comments are misclassified as "general_info" intent class.

5.2 Future work

In the future, more data will be collected for existing intent categories and for new intent categories. Besides, data will be created in the form of balanced data. There is also a plan to experiment with the combination of other pretrained model and deep learning model to improve the performance of the deep learning model. Fine-tuning is a process that takes a model that has been already trained is re-trained language model in Natural Language Processing (NLP). As a result of fine-tuning procedure, there is a plan to use fine-tuning procedure is used in next experiments.

AUTHOR'S PUBLICATIONS

[1] Ei Myat Myat Noe, Hsu Myay Mo, "Intent Classification of User Comments in Myanmar Language on Social Media Shopping Pages", National Journal of Parallel & Soft Computing, Yangon, Myanmar, 2022.

APPENDIX

As shown in table 5, the two methods perform with different advantages and disadvantages. The intent names have been altered to be more understandable and users readable.

No	User Input Text (Actual data)	Predicted Intent	
		Only Word Segmentation	**Word Segmentation and NER**
1	ကင်မရာထောက်တိုင် ဝယ် လို့ ရလား	purchase	purchase
2	တာချီလိတ် အိမ်ရောက် ငွေချေ ရလားဗျ	payment	job
3	အလုပ် လိုချင် ပါတယ်	job	job
4	ဆက်သွယ် လို့ ရမဲ့ ဖုန်းနံပါတ်လေး ပေးပါလား	contact_info	contact_info
5	ပဲခူး ပို့ပေး လားဗျ	delivery	contact_info
6	ထရိုင်ပို့စတန်း က အတိုးအလျှော့ရလား	general_info	purchase
7	ကျေးဇူးပါ အရမ်း သုံးလို့ ကောင်းပါတယ်	opinion	opinion

Table 5. Difference between two methods

REFERENCES

[1] B. Jang, M. Kim, Gaspard H., S. Kang, J. W. Kim, "Bi-LSTM Model to Increase Accuracy in Text Classification: Combining Word2vec CNN and Attention Mechanism", Appl. Sci. 2020, 10, 5841; doi: 10.3390/app10175841.

[2] D. Choi, I. Park, M. Shin, E. Kim, D. R. Shin, "Korean Erroneous Sentence Classification With Integrated Eojeol Embedding", Digital Object Identifier 10.1109/ACCESS.2021.3085864, June 3, 2021.

[3] D. Chen, C. Yin, "Data Augmentation for Intent Classification", 35th Conference on Neural Information Processing Systems (NeurIPS 2021), Sydney, Australia.

[4] H. Purohit, G. Dong,V. Shalin, K. Thirunarayan, A. Sheth, "Intent Classification of Short-Text on Social Media," 2015 IEEE International Conference on Smart City/SocialCom/ SustainCom (SmartCity).

[5] H.M. Mo, and K.M. Soe, "Syllable-Based Neural Named Entity Recognition For Myanmar Language", International Journal on Natural Language Computing (IJNLC), Vol.8, No.1, February 2019.

[6] H.M. Mo, and K.M. Soe, "Myanmar named entity corpus and its use in syllable-based neural named entity recognition ", International Journal of Electrical and Computer Engineering (IJECE), Vol.10, N0.2 pp. 1544~1551, April, 2020.

[7] H. N. Yeom, M. Hwang, M. N. Hwang, H. Jung, "Study of Machine Learning Classifier and Feature Set Selection for Intent Classification of Korean Tweets about Food Safety", J. of infosci. theory and practice 2(3): 29-39, 2014 http://dx.doi.org/10.1633/JISTaP.2014.2.3.3.

[8] J. Zhang, Y. Ye,Y. Zhang, Likun. Q., B. Fu, Y. Li, Z. Yang, J. Sun, "Multi-Point Semantic Representation for Intent Classification", The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20), 2020.

[9] J. Lim, S. Son, S. Lee, C. Chun, S. Park, Y. Hur, H. Lim, "Intent Classification and Slot Filling Model for In-Vehicle Services in Korean", Applied Sciences. 2022; 12(23):12438. https://doi.org/10.3390/app122312438, 5 December 2022.

[10] K. Yawata, T. Suzuki, K. Kiryu, K. Mohri, "Performance Evaluation of Japanese BERT Model for Intent Classification Using a Chatbot", The 35th Annual Conference of the Japanese Society for Artificial Intelligence, 2021.

[11] L. Pan, C. W. Hang, A. Sil, S. Potdar, "Improved Text Classification via Contrastive Adversarial Training", the Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-22), 2022.

[12] M. Kroll and M. Strohmaier, "Analyzing human intentions in natural language text," in Proceedings of the fifth international conference on Knowledge capture. ACM, 2009, pp. 197–198.

[13] M. Roman, A. Shahid, S. Khan, A. Koubaa, L. Yu, "Citation Intent Classification using Word Embedding", the State Key Laboratory of Computer Architecture (ICT, CAS) Open Project under Grant CARCHB202019, vol 9, 2022.

[14] N. Chen, X. Su, T. Liu, Q. Hao, M. Wei, "A benchmark dataset and case study for Chinese medical question intent classification", BMC Med Inform Decis Mak 20 (Suppl 3), 125 (2020).

[15] Purohit, H., Dong, G., Shalin, V. L., Thirunarayan, K., & Sheth, A. P. "Intent Classification of Short-Text on Social Media", The Ohio Center of Excellence in Knowledge-Enabled Computing (kno.e.sis), 2015.

[16] P. Zhang, L. Huang, "Research on Chinese Intent Recognition Based on BERT pre-trained model", ICMAI 2020: Proceedings of the 2020 5th International Conference on Mathematics and Artificial Intelligence, 29 May 2020.

[17] Q. Chen, Z. Zhuo, W. Wang, "BERT for Joint Intent Classification and Slot Filling", arXiv: 1902.10909v1 [cs.CL], 28 Feb 2019.

[18] Simonyan, K. and Zisserman, A. (2015).Very Deep Convolutional Networks for Large-Scale Image Recognition.

[19] S. Wang and C. D. Manning, “Baselines and bigrams: Simple, good sentiment and topic classification,” in ACL, 2012, pp. 90–94.

[20] S. Rizou, A. Paflioti, A. Theofilatos, A. Vakali, G. Sarigiannidis, K. Ch. Chatzisavvas, “Multilingual Name Entity Recognition and Intent Classification Employing Deep Learning Architectures”, Simulation Modelling Practice and Theory, Volume 120, November 2022, 102620.

[21] T. N. Tun, K.M. Soe, “Intent Classification on Myanmar Social Media Data in Telecommunication Domain Using Convolutional Neural Network and Word2Vec”, 2020 23rd Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)

[22] Z. M. Kim, Y. S. Jeong, J. Hyeon, H. Oh, and H. J. Choi,” Classifying Travel-related Intents in Textual Data ,” Int'l Journal of Computing, Communications & Instrumentation Engg. (IJCCIE) Vol. 3, Issue 1 (2016) ISSN 2349-1469 EISSN 2349-1477.