A Multi Class Classification to Detect Original Form of Kaomoji using Neural Network

Noriyuki Okumura[1][a] and Rei Okumura[2]
[1]*Faculty of Modern Social Studies, Otemae University, Japan*
[2]*Advanced Course of Mechanical and Electronic System Engineering, National Institute of Technology, Akashi College, Japan*

Keywords: Kaomoji, Original Form, Neural Network, Middle Layer.

Abstract: In this paper, we propose a multi-class classification method for Kaomoji using feed forward neural network. Neural network has some units in each layer, but the suitable number of units is not clear. This research investigated the relation between the number of units and the accuracy of multi-class classification method.

1 INTRODUCTION

In this paper, we report on the estimation of the original form of Kaomoji, emoticons of Japanese style, which is one of the classification tasks of Kaomoji. Kaomojis are represented not only by ASCII characters that are half-width characters but also by combinations of various characters, including full-width characters such as Japanese characters (Kanji, Hiragana, and Katakana). At present, the total number of Kaomoji has already exceeded over 100,000. Besides, it is also possible to compose Kaomoji of shapes that do not fit on one line by composing Unicode. According to Kazama et al., it is possible to extract millions of face Kaomoji by using Twitter logs (Kazama et al., 2016). On the other hand, any researchers did not establish a framework for comprehensively treating Kaomoji, and each researcher has only a large-scale dictionary of Kaomoji.

In this paper, in order to put together a wide variety of Kaomoji, we aim to define the original form of Kaomoji and estimate the original form from arbitrary Kaomoji. We have already proposed a method using neural networks as a method for estimating Kaomoji, but we did not verify the validity of the number of units to be used in the middle layer. Therefore, in consideration of the possibility that certain Kaomoji, especially theatrical type Kaomoji, have multiple original forms, the number of units in the middle layer in multiclass classification is investigated.

As a result of the research experiment, we confirmed that by preparing 6,500 units in the case of the middle layer, we could obtain the best results in terms of learning time and accuracy rate.

[a] https://orcid.org/0000-0003-1149-1645

2 RELATED WORK

Bedrick et al. try to detect Kaomoji using PCFG (Probabilistic context-free grammar). (Bedrick et al., 2012) The target of extracting Kaomoji is tweets posted on Twitter. This method only uses the rules of PCFG (Probabilistic context-free grammar) and does not define the original form of Kaomoji.

Kazama et al. proposed Kaomoji detection algorithm. (Kazama et al., 2016)(in Japanese) their method is to analyze articles posted on SNS such as Twitter and extract Kaomoji-like strings. By the method of Kazama et al., it is possible to extract character strings that can be regarded as a large number of Kaomoji as similar as Bedrick et al.. Their method, however, only rules the composition of Kaomoji and is not suitable for grouping Kaomoji as in this paper.

Ptaszynski et al. proposed CAO system as a framework for the comprehensive treatment of Kaomoji. (Ptaszynski et al., 2010) The CAO system not only extracts the character string that can be regarded as Kaomoji but also aims to extract the emotion that the extracted Kaomoji expresses. In particular, this method extracts Kaomoji using the eye-mouth-eye sequence (triplet) as the basis of Kaomoji as a feature of the Kaomoji. This method is similar to this research in that it defines a basic string representing Kaomoji. As we

Okumura, N. and Okumura, R.
A Multi Class Classification to Detect Original Form of Kaomoji using Neural Network.
DOI: 10.5220/0008366203770382
In *Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2019)*, pages 377-382
ISBN: 978-989-758-382-7

have not made detailed definitions such as (Okumura, 2016) for considering symmetry, it is not suitable for grouping.

The author's team is developing a classification method for Kaomoji using neural networks and cosine similality. (Okumura and Okumura, 2018)(Okumura, 2017) As a result, it is possible to extract the only original form inferred from any Kaomoji with about 70 % accuracy rate. On the other hand, we do not implement a multiclass classification for Kaomoji with multiple primitives.

In this paper, we aim to construct a system that can be estimated even if it belongs to multiple classes by labeling a string that can be regarded as Kaomoji as a class called an original form. Also, in this paper, we examine the number of units in the middle layer of the neural network, which is necessary to estimate the original form from the character string that can be regarded as a Kaomoji.

3 THE METHOD TO ESTIMATE THE ORIGINAL FORM OF KAOMOJI CORRESPONDING TO MULTI-CLASS CLASSIFICATION

Our conventional method aims at outputting one of 3,110 original form of Kaomoji. However, for example,
(￣▽￣)ノ" ○(*　　*)○ヾ(￣▽￣) ヨチヨチ
in the case of such Kaomoji, the system cannot judge which original form to extract because of fo including three types of character sequences considered to represent a face. Therefore, extracting just one original form is not enough as a grouping, and it is necessary to identify all the faces included in the Kaomoji-like character sequence and its original form.

In this paper, we implement multiclass classification using fixed-length input feed-forward neural network using the character Embedding. In the previous example, we have to construct a model that is correct if we can extract the three original forms ((￣▽￣), (　　), (￣▽￣), strictly, (￣▽￣) has appeared twice, so our system have to extract two types of original form of Kaomoji).

3.1 Multiclass Classification of Kaomoji

The neural network used in this paper is a simple model with only one middle layer. On the other hand, in the case of Kaomoji with linguistic features, it is known that the information to be given to the input layer is insufficient with the One-hot vector. Therefore, this system vectorizes the input (Kaomoji) using the character Embedding. At the time of writing this article, the longest of the emoticons registered in the emoticon database collected is 65 characters, so the input to the Embedding layer is 65 units. Each character input to the Embedding layer is converted to a 100-dimensional vector in order from the left side of the emoticon and combined in order. For Kaomoji less than 65 characters, generate a fixed-length input vector with zero paddings (NULL characters) for the shortfall. The figure 1 shows the configuration of the neural network adopted in this paper.

In this paper, we change the number of units in the middle layer in the figure 1, and we want to derive the appropriate number of units based on both the evaluation by the correct answer rate (described later) and the learning time. Some researchers noted that although there is an argument that there is a standard of (number of input units + number of output units) x 2/3 as a standard of the number of units, it is various factors such as the complexity of the problem to be solved and the size of learning data. Because of the influence, it does not go beyond the range of heuristics.

3.2 Evaluation Method

In our past studies, it was regarded as correct as long as at least one correct original form is included in the outputted original form group as an evaluation scale in the estimation of the emoticon base form. However, in multiclass classification, we have to evaluate whether outputs of our system include all the original forms. In this paper, if our system classified three original forms, then; the conventional evaluation method (Easy) is corresponding to the correct if one of the output is a correct answer; Normal evaluation method is corresponding to the ratio between correct answers in output and the number of correct answers (Normal); Hard evaluation method is corresponding to the correct if the system's output is corresponding to all of the correct answers.

For example, if there are three types of correct answers and the estimated results are these, For example, if there are three types of correct answers and the estimated results are these, then the evaluation method for Easy answers correct, the evaluation method for Normal calculates a ratio as a correct rate of 2/3, and the evaluation method for Hard answers incorrect. Evaluation is performed by the following formula 1, 2, 3.

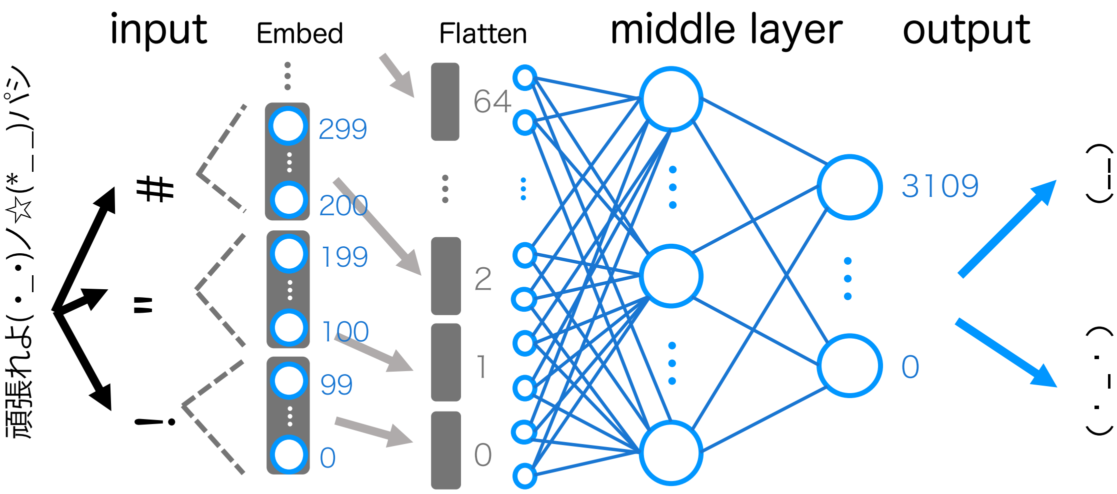

Figura 1: The Model to estimate the original form of Kaomoji.

$$Eval_{Easy} = \frac{E}{N} \quad (1)$$

$$Eval_{Normal} = \frac{1}{N}\sum\frac{Normal}{NormalCorrect} \quad (2)$$

$$Eval_{Hard} = \frac{H}{N} \quad (3)$$

E. E is corresponding to the number that the output has one of the original forms at least

Normal. Normal is corresponding to the number of correct answers in output

NormalCorrect. NormalCorrect is corresponding to the number of correct answers for each Kaomoji

H. H is corresponding to the number that all output equals the original forms

N. the number of test data

In $Eval_{Easy}$, the ratio of the number of correct answers to all evaluation targets is calculated, in $Eval_{Norm}$, the average accuracy rate of all evaluation targets is calculated, and in $Eval_{Hard}$, $Eval_{Easy} Calculatethesameratioas$, and investigate the change in accuracy rate depending on the number of units.

In this evaluation, the average accuracy rate of this paper is calculated under 10-fold cross validation using 28,296 pairs of Kaomoji-like character strings and original forms. In the following sections, we examine the transition of the average accuracy rate of 10-fold cross-validation and the time for learning.

4 RESULT

In this section, we describe the results of evaluating neural network method in figure 1 by the method described in section 3.2. The minimum number of units in the middle layer is 500, and the results show the tendency in increments of 500 up to 10,000.

4.1 Evaluation for Loss

Figure 2, figure 3 show the transition of loss in learning data and evaluation data. The vertical axis shows the output from the loss function, and the horizontal axis shows the Epoch number.

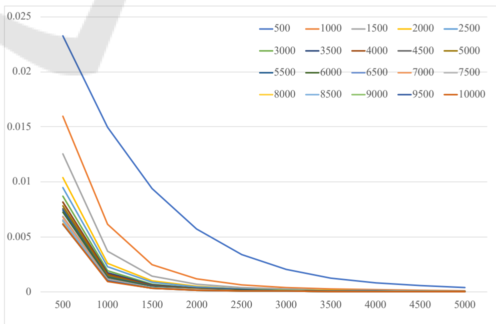

Figura 2: Loss of training data.

4.2 Evaluation for Training Data

Figure 4, figure 5, figure 6 the transition of the evaluation value by the method. The vertical axis indicates the accuracy rate, and the horizontal axis indicates the number of Epochs.

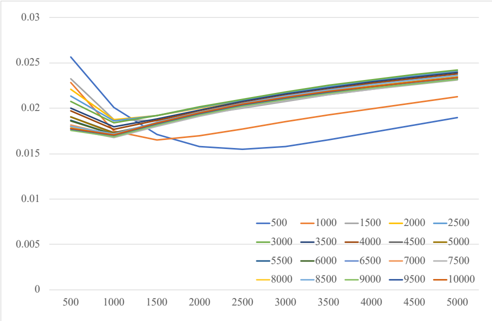

Figura 3: Loss of evaluation data.

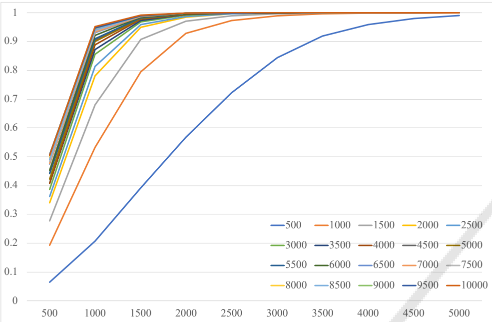

Figura 4: Evaluation method "Easy" for training data.

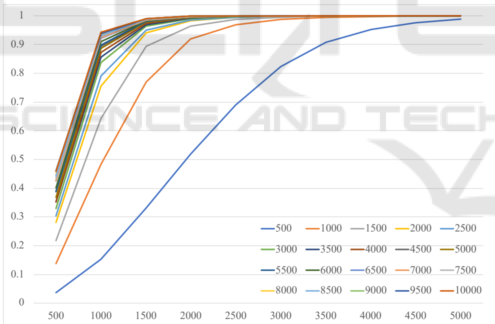

Figura 5: Evaluation method "Normal" for training data.

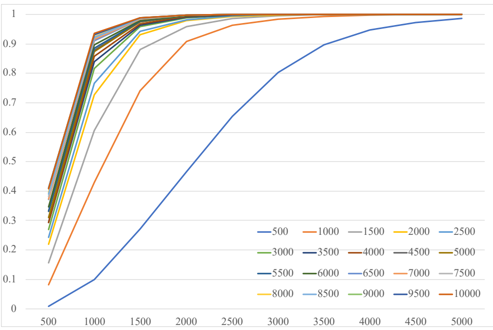

Figura 6: Evaluation method "Hard" for training data.

4.3 Evaluation for Evaluation Data

Figure 7, figure 8, figure 9 the transition of the evaluation value by the method. The vertical axis indicates the accuracy rate, and the horizontal axis indicates the number of Epochs.

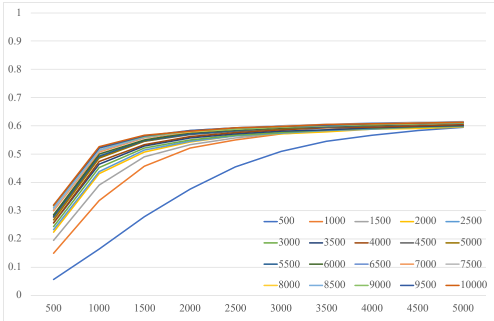

Figura 7: Evaluation method "Easy" for training data.

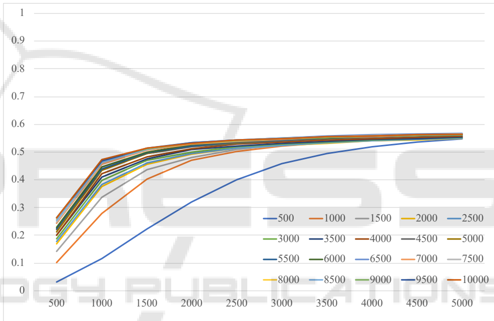

Figura 8: Evaluation method "Normal" for training data.

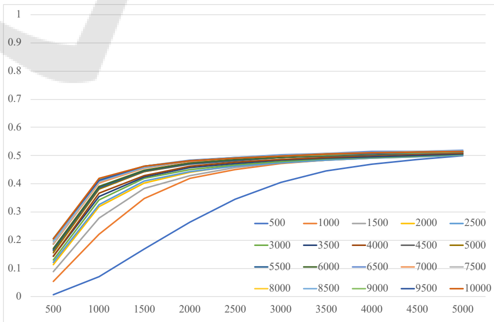

Figura 9: Evaluation method "Hard" for training data.

4.4 Time for Evaluation

Figure 10 shows the time for each evaluation. The vertical axis shows the required time.

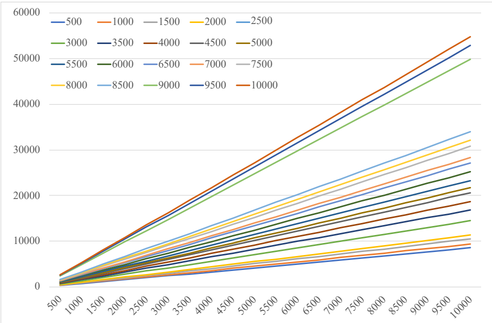

Figura 10: Time for 10-fold cross-varidation.

5 DISCUSSION

Figure 2, figure 3 in section 4.1 show that the training process is advancing rapidly up to around 1000 Epoch. Moreover, the loss in evaluation data increases after 1000 Epoch. Therefore, it is sufficient for our neural network model in the proposed model that the number of training Epoch is about 1,000 Epoch. There is no difference depending on the number of units except 500 units and 1,000 units.

According to each evaluation (Easy, Normal Hard) shown in section 4.2, in all evaluations, the evaluation values show the same tendency except in the case of 500, 1,000, and 1,500 units. The correctness rate is 90% or more for training data by preparing the middle layer of at least 2,000 units or more from the evaluation value of around 1000Epoch. Therefore, learning converges at around 1000 Epoch.

According to each evaluation (Easy, Normal, Hard) shown in section 4.3, although the evaluation rate for Easy is only about 10% high, there is no difference in the evaluation values for Normal and Hard. From this, when performing multiclass classification, it is rarely output more than the total number of original forms contained in emoticon-like character strings. The number of primitives output by estimation is equal to the total number of primitives (number of correct primitives) included in the Kaomoji-like character sequence. On the other hand, the accuracy rate is about 50%. Therefore, the future task is to improve performance. The correct answer rate around 1,000 Epoch is the highest with an accuracy of about 42% for 10,000 units, but the number of units calculated based on the rule of thumb described in section 3.1 is 6500 (= The accuracy rate in the case of (6500 + 3110) x 2/3 = 6406 = 6500) is about 41%, which is a small difference. Also, the overall tendency is that the accuracy rate slightly increases as the number of units increases, but the accuracy rate may be lower than in the case of the number of units according to the rule of thumb, so increasing the number of units is not better.

Finally, we will consider the time required for learning shown in section 4.4. The time required for learning increases as the number of units increases, as shown in figure 10.

A smaller number of units is more effective than a more significant number of units from the viewpoint of the required time for 10-fold cross-validation. On the other hand, since it is difficult to improve the accuracy rate, it is considered better to adopt the number of units near the number of units based on heuristics. Besides, when the number of units is 9,000 or more, the required time tends to increase rapidly, so the number of units less than 8,000 is considered appropriate for the model adopted in this paper.

6 CONCLUSIONS

In this paper, we investigated the number of units in the middle layer in a feed-forward neural network to estimate the original form of Kaomoji. We confirmed experimentally the optimum value of the number of units based on the empirical rules and found that it is a model that does not deviate from the empirical rules. When the number of units in the middle layer is 6,500, we confirmed that training exceeds 3,500 Epoch and over 50% in any evaluation method, but the problem occurs that it takes too much time for training.

As future work, the proposed system is based on the character Embedding as information to be given to the input layer, but systems can deal with only the characters that appear in the database of Kaomoji, so it is necessary to consider the Embedding method that can correspond to all Japanese characters. Also, since the input has a fixed length of 6,400, it is necessary to apply a model compatible with variable-length input such as a recurrent neural network and extend it to a system compatible with an emoticon of any length. There is. Similarly, in this paper, the middle layer is considered to be one layer, but we will investigate the accuracy rate in the case of multiple layers, and we would like to clarify the relationship between the number of layers and the number of units as a model used for the analysis of Kaomoji.

ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENSHI Grant Number 18K11455.

REFERENCES

Bedrick, S., Beckley, R., Roark, B., and Sproat, R. (2012). Robust kaomoji detection in twitter. In *Proceedings of the Second Workshop on Language in Social Media*, pages 56–64. Association for Computational Linguistics.

Kazama, K., Mizuki, S., and Sakaki, T. (2016). Study of sentiment analysis using emoticons on twitter(in japanese). *The 30th Annual Conference of the Japanese Society for Artificial Intelligence*, JSAI2016:3H3OS17a4–3H3OS17a4.

Okumura, N. (2016). 顔文字の原形抽出 (in japanese). *Proceedings of the 22nd annual meetingof the Association for Natural Language Processing*, NLP2016:P1–1.

Okumura, N. (2017). A large scale knowledge base representing the base form of kaomoji. In *Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management - Volume 2: KEOD,*, pages 246–252. INSTICC, SciTePress.

Okumura, R. and Okumura, N. (2018). A method to estimate the original form of kaomoji using neural network with character embedding function(in japanese). *The 32nd Annual Conference of the Japanese Society for Artificial Intelligence*, JSAI2018:2B405–2B405.

Ptaszynski, M., Araki, K., Dybala, P., Rzepka, R., and Maciejewski, J. (2010). Cao: A fully automatic emoticon analysis system based on theory of kinesics. *IEEE Transactions on Affective Computing*, 1:46–59.