

A Conversational System with Enhanced Emotion Expression by Using Emoji

Wang, Ziwei

(出版者 / Publisher)

法政大学大学院情報科学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 情報科学研究科編

(巻 / Volume)

15

(開始ページ / Start Page)

1

(終了ページ / End Page)

6

(発行年 / Year)

2020-03-24

(URL)

<https://doi.org/10.15002/00022715>

A Conversational System with Enhanced Emotion Expression by Using Emoji

Wang Ziwei
Graduate School of Computer and Information Sciences
Hosei University
Tokyo, Japan
ziwei.wang.8e@stu.hosei.ac.jp

Abstract—Wechat, a Chinese messaging app, being used by over a billion users monthly. Messaging apps like Wechat use robot systems to predictively generate accurate conversational responses for users. However, users often prefer to convey the emotional expression necessary for human communication in pictorial form with Emoji. In this paper I have used the existing system to generate responses primarily with sentimental analysis, with which the primary text-based reply is fused with emoji stickers so as to be able to convey a more accurate emotional response. This was accomplished through carrying out an emotional analysis of the input text which is then used to select the appropriate emoji stickers to make a response. This analysis also accounts for age and emotional consistency in selecting the most appropriate emoji to convey emotion. As a result, the system is enhanced by facilitating richer communication between users, who are able to convey their actual emotions more swiftly. In this paper, some conversation cases are run using the enhanced system and compared with the existing original system.

Keywords—WeChat; Emoji; Sentiment Analysis

I. INTRODUCTION

With the development of Internet technology, the so-called "Internet era", has been ushered in, especially the mobile Internet characterized by intelligent mobile devices. In addition to the rapid development of artificial intelligence technology in recent years, chat dialogue systems have become commonplace in all walks of life. Microsoft has launched Xiao Bing, a chat robot based on social computing, Baidu has launched a chat robot for interactive search, and the production of chat robots has been realized. Apple's Siri personified Q & A system uses Siri to read text information, set alarm clocks, query restaurant descriptions, and ask about the weather, and so on.

This can provide attractive access to information, greatly facilitate people's lives, and, to a certain extent, meet the needs of chat participants. The popularity of smartphones has led to the popularity of WeChat, a smartphone end-to-end communication software, with more than 1 billion monthly active users in 2018. WeChat has widely influenced people's daily life. Maoyuan Shen has implemented a chat robot on the WeChat platform which can read data, such as the conversation text of a chat counterpart in WeChat. The robot acts as a friend and can interact with WeChat users in the form of automatic response generation [1].

Emotions, such as joy, anger, worry, thinking, sadness, fear, shock and so on, are an important factor in understanding and expressing information. At present, research work related to chat robots is trying to enable a chat robot to generate appropriate dialogue or replies with only text-based content, but without visually-based emotion expressions like humans do in face-to-face communication. In common SNS chat systems like LINE, WeChat, and Facebook, people often use emoji as well as text to express various feelings and emotions in their daily SNS conversions. Emoji [2], as one of the most popular and convenient conveyors of emotion expression, make dialogue and communication livelier and smoother, so they are widely used by people of all ages. Some of the common emoji expressions have become almost completely accepted as conventional meanings of the emotion which they are intended to express.

In order to solve the problem of the lack of visual-based emotion expression that existing chat robot systems have in generated conversations, this paper is focused on adding emotion expression with emoji to an existing SNS WeChat conversational system.

II. THE EXISTING CONVERSATIONAL SYSTEM

The existing conversational system developed by Maoyuan was based on a word2vec model and sentence embedding Seq2Seq [3] with an attention model. In the Word2vec [4] model, text is transformed into computable vectors. The seq2seq model adopted an LSTM network in which word2vec was embedded and trained using an existing conversation corpus. The trained model was able to predict or generate reply sentences. This system is basically a text-based reply sentence generation system, and its overview is given below in Fig. 1.

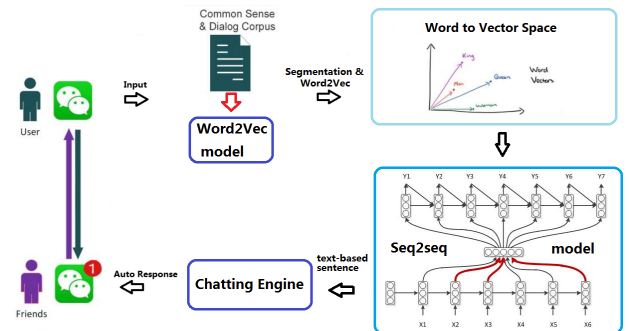


Fig. 1. Maoyuan's conversational system

As it can be seen that the output from the word2vec is the input for the Seq2Seq & Attention model. The output of seq2seq model with word attention is a reply sentence which can be generated automatically. The examples are listed in the Table. 1. In order to make it easy to understand, I have translated the Chinese text into English.

TABLE I. Some examples of the results.

Input text	Reply text
You're so beautiful!	That's necessary.
Be modest, huh?	I was wrong!
It's so smug.	Yeah, I like it. You're afraid to say.
I like you.	I don't like you.
Do you think I'm cute?	Very! Not cute.
The flowers you sent are so beautiful.	Master, don't be sad. You still have me.

From the above automatic example results from the reply generation model, it can be seen that Maoyuan's conversational system can generate more or less reasonable replies to correspond to the content of its counterpart in most cases. However, in daily expression and communication, people do not only rely on a single language item or text. Moreover, in communication with others, people often need to convey certain feelings. Facial expressions, tone, gesture and other elements are often used, so as to facilitate the transmission of emotion. to carry out the process of communication successfully.

Therefore, at this level, Maoyuan's conversational system only produces a response to the text which is somewhat rigid, in which there is a lack of emotion expression. At the same time, when people use WeChat chat, they often choose to use emojis as a way to express their feelings. Emojis are small pictures that simulate a person's facial expressions, so they can convey emotions simply, intuitively and naturally. Therefore, if the appropriate emoji can be added to the reply text, the rigid output generated by Maoyuan's conversational system can be improved upon and an emotion output can be generated. For example, the emoji sticker of a smile depicts a smiling face to express happy emotions.

In addition, there are some responses that are not appropriate in terms of emotion expression. For example, input which says,

"The flowers you sent are too beautiful"

however, the reply generated is

"Don't be sad, you still have me"

Where, the speaker is obviously very happy, but the content of the reply is understood to convey a very sad mood and to give comfort. If an emotion analysis is performed on both sides of the conversation, it can guide the conversation system to generate appropriate responses according to the input's emotion classification. Therefore, the generated reply will have a better effect because the appropriate expression of feelings is added into the reply.

III. THE ENHANCED SYSTEM WITH EMOJI EXPRESSION

In order to improve the Maoyuan's conversational system mentioned above, so that it can produce a more vivid and appropriate reply, an enhanced conversational system with Emoji expression is proposed.

The system diagram is given below in Fig. 2.

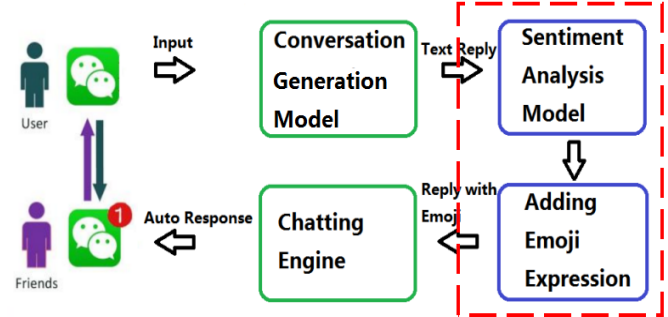


Fig. 2. The enhanced conversational system with emoji expression.

The part of the dotted line is newly added to the Maoyuan system, comprising two parts, namely, the emotion classification and the improved conversation system with emoji expression.

After a reply sentence is generated from Maoyuan's model, it is passed into sentiment analysis model for emotion classification. Then, according to the results of the sentiment analysis, the reply is improved with an emoji that can express the appropriate emotion. Finally, the improved reply is passed back to the chatting engine. The functions of these two parts are described below.

A. Regarding Sentiment Analysis

Emotion analysis refers to the classification of the text into two or more types of praise or derogatory meanings according to the meaning and emotion information expressed in the text. It is the division of the tendency, viewpoint and attitude of the author of the text.

Some researchers have been done into this issue and many classification methods have been tested [5]. Generally, these can be divided into machine learning methods and emotion dictionary methods, and after comparison of several methods, and finally an emotion dictionary method was used to carry out a follow-up experiment.

B. Improvement with Emoji Expression

1) Using Emoji Expression

According to the result of the emotion classification model to reply, the corresponding Emoji Expression. is used at the end of reply. However, this alone is not enough, according to the normal habit of conversation, people will use some expressions to replace or supplement the text to express certain feelings, such as: *"you are great, such a difficult problem can be solved!"*

A smiling face pattern can be added here to express the positive emotion of praise. Therefore, by comparing the feature word vector, and placing the Emoji Expression of the emotion dictionary to the middle or front of the reply sentence, the improved reply is more vivid.

2) Sentiment Consistence Analysis

Sentiment consistence analysis is the analysis of the emotion tendency for both input text and reply text, respectively. In the case of an emotion tendency mismatching the input and reply, it is necessary to improve the emotion of the reply sentence in order to ensure a stable and smooth input-and-reply dialogue. Next, the method of emotion classification and the method of applying emotion classification to improve the conversation system will be introduced in detail.

IV. SENTIMENTAL ANALYSIS

The emotion recognition task is regarded as a classification problem, and two classification models have been constructed. One is the emotion tendency classification model based on SVM, the other is LSTM (long and short memory neural network) model. In addition to constructing two classification models, this paper also studies an analysis method based on an emotion dictionary, that is, using the emotion tendency and intensity information of words or phrases recorded in the emotion dictionary to classify the text.

A. SVM

The Support Vector Machine is a class of generalized linear classifiers for binary classification of data, according to supervised learning. It is considered to be a better method for text classification. The implementation flowchart of a SVM model is giving in Fig.3. The corpus file is loaded with emotion tags and the data is imported. This paper uses Weibo with emotion markers, a shopping after-sales review corpus for training, including 10,679 positive corpus items, 10,428 negative corpus items. Based on the Word2vec tool of deep learning, we take the mean value of all the word vectors of each sentence and train the low-dimensional word vectors which contain deep semantic information.

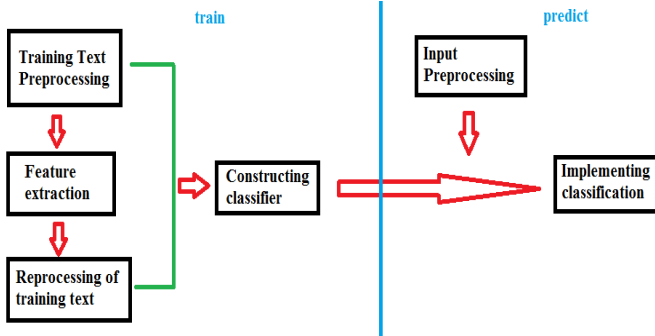


Fig. 3. The flowchart of SVM model.

The training model is implemented in the Python language, and the training process is implemented by a svm function in a sklearn library, using a sklearn.svm.SVC for training. Firstly, the kernel function of the training is set, then the trained model is trained and saved, and then the word vector of the input is obtained according to the low-dimensional word vector model generated by Word2vec. Finally, according to the SVM classifier model trained by SVC, the emotion classification value of sentences is predicted. The partial of experimental results for predicting one's conversation text's emotion tendency as positive and negative is shown in the Fig. 4.

```

还不错，这个可以正常使用 positive
/home/shenmaoyuan/anaconda3/lib/python3.6/site-packages/sklearn/bi
C from version pre-0.18 when using version 0.19.1. This might lea
isk.
UserWarning)
外观漂亮，我喜欢这个。 positive
/home/shenmaoyuan/anaconda3/lib/python3.6/site-packages/sklearn/bi
C from version pre-0.18 when using version 0.19.1. This might lea
isk.
UserWarning)
买错了，只能勉强凑合着用。 negative
/home/shenmaoyuan/anaconda3/lib/python3.6/site-packages/sklearn/bi
C from version pre-0.18 when using version 0.19.1. This might lea
isk.
UserWarning)
这里设施不算新，但是挺温馨的，卫生也好 positive
/home/shenmaoyuan/anaconda3/lib/python3.6/site-packages/sklearn/bi
C from version pre-0.18 when using version 0.19.1. This might lea
isk.

```

Fig. 4. SVM predicted results.

B. LSTM

We built an LSTM model that takes as input word sequences. This model is able to take word ordering into account. We use pre-trained word embeddings to represent words, feed them into an LSTM to predict the most appropriate emoji [6].

The steps of LSTM sentiment analysis are as follows: the first step is loading the corpus. This experiment uses the same corpus as the SVM. The layers are then embedded: Training a Word2vec model creates a word dictionary and returns the index of each word, the word vector, and the word index for each sentence. Finally, the model is trained and saved. The accuracy score of the trained model was 0.918.

According to the model trained by the LSTM, the emotion classification value of sentences is predicted. The predicted results are shown in Fig. 5.

```

loading model.....
loading weights.....
今天的天气真的很不错。 positive
loading model.....
loading weights.....
房子漏水了，讨厌死了。 negative
loading model.....
loading weights.....
和想象中差好多，不好！ negative
loading model.....
loading weights.....
还不错，这个可以正常使用 positive

```

Fig. 5. LSTM predicted results2.

C. Emotion Dictionary

The emotion dictionary method achieves an emotion score through a series of emotion dictionaries and rules. The operation steps are shown in the Fig. 6.



Fig. 6. The flowchart of emotion dictionary model.

The first is text preprocessing. This paper uses the corpus provided by Hownet, Taiwan University NTUSD, Tsinghua University-Li Jun Chinese commendatory Dictionary and other corpora to construct the experimental emotion dictionary. Then the jieba.cut word segmentation tool is used to segment the

Chinese text, and then remove the stop words. The most important part is to calculate the total score of emotion tendency. The flow chart of this part is shown in the Fig. 7.

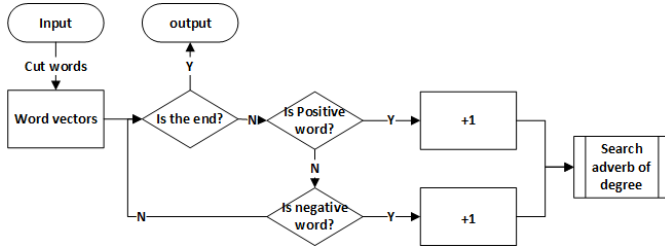


Fig. 7. The flowchart of calculating the sentiment score.

Judgement as to whether positive or negative evaluative words exist or not is done word-by-word. If the positive evaluative words exist, the positive score will add 1. If the negative evaluative words exist, the positive score will add 1. Then there is a check to see if degree adverbs exist, and the cumulative score will be multiplied by its weight. The judgment and weight of degree adverbs are shown in the Fig. 8.

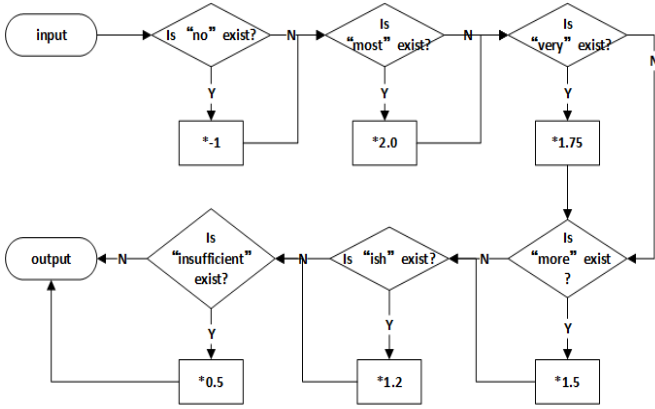


Fig. 8. The weight of adverbs of degree

The implementation result of the emotion dictionary is shown in the Fig. 9.

今天的天气真的很不错。 1.8
房子漏水了，讨厌死了。 -2
和想象中差不多，不好！ -3
还不错，这个可以正常使用 2.5
外观漂亮，我喜欢这个。 2
买错了，只能勉强凑合着用。 -1
这里设施不算新，但是挺温馨的，卫生也好 2.2

Fig. 9. The result of adverbs of emotion dictionary model

Comparing the results of the three emotion analyses, it is possible to see that the training corpus is not large enough, so the accuracy of the two modeling methods is not as good as that of the emotion dictionary. Especially in the presence of attributives and other special cases, the effect is not good. Moreover, these two modeling methods can only distinguish between positive and negative emotion pairs, while the emotion dictionary can get emotion values, which are more conducive to matching emoji and emotion pairs with finer granularity. Therefore, this experiment selects the optimal emotion dictionary method to carry on the emotion analysis.

V. GENERATED RESPONSE BY ADDING EMOJI EXPRESSION

A. Adding Emoji Expression

The 10 common emojis used in this paper are often used by people in everyday conversation to express strong emotions [7]. In his research, he collected a large number of chat records, including a large number of emoticons, and based his statistics on the frequency with which 88 expressions are used on WeChat. A few emoticons are widely used, while most emoticons are used less frequently. In this paper, the 10 expressions that are both the most frequently used and express emotion at the same time were chosen.

In order to achieve emotion conversational, we should add an expression that can express the appropriate feelings in the generation of the response. The steps are as follows:

a) Make a emoji sticker and the its emotion category value;

In order to find the correspondence between emoji and emotion values, 100 sentences were taken to calculate the emotion values. The results are shown in Fig. 10.

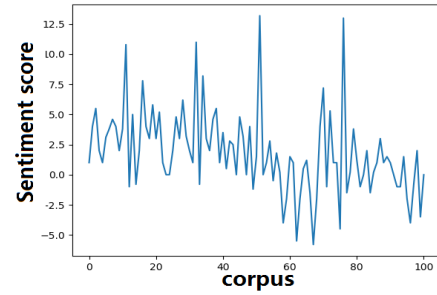


Fig. 10. The emotion values of 100 sentences.

It can be seen from the Fig. 10, showing the emotion values of 100 sentences, the values of most sentences are concentrated in the $[-5, 5]$ range, so the definition of the emotion value and the corresponding expression are shown in Table.2.

TABLE II Emotion value and emoji

Emotion value	Emoji	The depiction of Emoji
score>8	😄	excited expression
5<=Score<8	😂	laugh
2.5<=Score<5	😊	happy expression
1<=Score<2.5	😏	complacent expression
0<Score<1	🙂	smile
-1<=Score<0	😬	worried expression
-2.5<=Score<-1	😞	aggrieved expression
-5<=Score<-2.5	😓	sad expression
-8<=Score<-5	😡	tormented expression
Score<-8	😱	break down

- b) Calculate the emotion score of the generated reply.
c) Produce the corresponding emotion expression.

B. Analysis of Sentiment Consistency

Because the factors of emotion expression are not taken into account, the emotion tendency of the reply sentences generated by Maoyuan's system is often inconsistent with or even conflict with input, so sometimes the reply is not coherent with the input sentence. Therefore, the participants' experience is not good. However, in people's real chat, there is often a continuous emotion tendency to run through the conversation. For example,

Zhang: *I got 99 points in the exam, not 100 points,*
Robot: *What a pity! Don't mind.*

In fact, what Zhang wanted to say was that he did well in the exam and wanted to share his joy with others, but the conversation generated by the robot was soothing and sad, so it seemed awkward and was unable to continue talking.

In order to improve this kind of input, we add contextual emotion judgment. When the emotion classification of inputs and replies are inconsistent, we improved the reply.

The process for this feature is as follows: (1) Judge the emotion of the input, write it down as S1; (2) Judge the emotion of the reply, write it down as S2; (3) If S1 and S2 are the same, add expressions according to S2's emotion classification and output them.

If S1 and S2 are different, the input is adjusted by adding words related to the emotion classification of the input, and then it is passed to the system again for guiding the model to predict the correct emotion classification reply.

VI. EXPERIMENTS AND RESULT ANALYSIS

A lot of experiments, up to 1000 groups, 80% of the dialogue achieved the desired results. At present, it mainly achieves better improvement in the following three aspects. The experimental results and some samples were listed.

A. Add Emojis to the Generated Sentence

Adding emoji corresponding to the emotion tendency of the reply. It makes human-computer conversation more vivid and vivid, rather than making people feel like talking with machines, as if they were communicating with people.

In Fig. 11, the input was "Hello" and the reply was "hello, and I hope you have a happy day." and a happy emoji added before the reply sentence. This emoji conveys happy feelings. User felt the conversation is real and vivid.



Fig. 11. Add emoji to the generated sentence (1)

In Fig. 12, the input was "Eat less junk food" and the reply was "Occasionally indulge, but dare not eat more" and a frown expressions are added at the end of the reply. User felt a sense of helpless. After adding emoji, user felt the reply has a helpless meaning.



Fig. 12. Add emoji to the generated sentence (2)

In Fig.13, the input was "Don't show off your love relationship." and the reply was "OK, you have known our love (This is blessed by you.)" and an expression of "admiration" in the middle of a reply sentence. This emoji conveys the feeling of blessing. Users feel praised, happy and willing to continue chatting.



Fig. 13. Add emoji to the generated sentence (3)

B. Analysis of Sentiment Consistency

Through machine learning, the enhanced system can distinguish the emotion of sentence, such as happiness, fear, anger, so as to better realize human-computer dialogue.

In Fig. 14, the input was "Give you a scare", but the reply generated by existing system is "I'm a little Lori. Uh-huh.". The reply is quite baffling and does not correspond to the input's sentiment. So we improve the reply to "Why scare me?". It's obviously more natural for user.



Fig. 14. Add emoji to the generated sentence (4)

In Fig. 15, the input text was "I'm angry today.", but the reply generated by existing system was "Happy Birthday!". The reply is very inappropriate. After improvement the new reply was "Who bullied you? I helped you". It's more natural and more like the mutual concern of friends.



Fig. 15. Add emoji to the generated sentence (5)

C. Add Different Expressions According to Age

Considering that interlocutors have different habits of using emoji at different ages, this paper makes a special adaptation to address this problem.

According to the report released by WeChat, "2018 WeChat Annual data report", people of different ages have different preferences for emojis [8], which is shown in Table. 3.

TABLE III Age group and the most common emoji stickers they use.

Age group	Emoji
the Post-70s	😏
the Post-80s	😁
the Post-90s	😂
the Post-00s	😭

So based on this idea, users are subdivided by age demographic. When talking to people of different ages, WeChat chat robots add emojis suitable for that age group to better express their feelings. It is more likely to cause emotional resonance with the interlocutor.

In the example shown in Fig. 16, if the interlocutor is a post-70s user, the most common expression is replaced with an emoji that indicates "sneering", adding a conversation item with such an expression. It is easier for the participants to have the same sense of intimacy as the conversation of his or her peers. The user experience effects of post-80s user and post-90s were shown as Fig. 17 and Fig. 18.

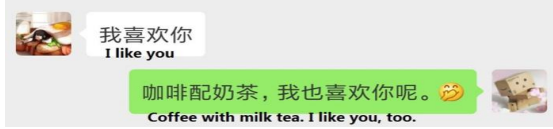


Fig. 16. Add emoji to the generated sentence (6)

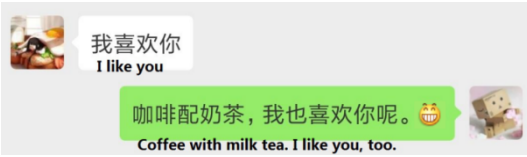


Fig. 17. Add emoji to the generated sentence (7)

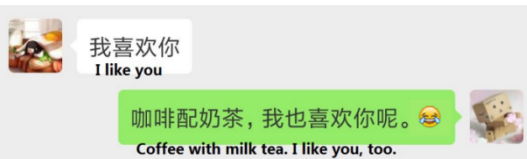


Fig. 18. Add emoji to the generated sentence (8)

Current research is only a beginning, in addition to the above three aspects, there are still some situations that have not been considered. Such as the analysis of rhetorical questions, the analysis of input as expression... These will be studied in the future.

Finally, in order to compare the enhanced system to the existing conversion system. A questionnaire was designed. 50 people were interviewed at different ages and occupations. In this questionnaire, 10 groups of the input texts and replies generated by these two systems were listed. Volunteers were asked to score the replies from 5 indicators: appropriate, emotion, vitality, continuity and comprehensibility. The following is to

add up the indicators of the results of the questionnaire and take the average. The results are shown in Fig. 19. Vertical axis represents the average of the total scores of 10 questions in 50 questionnaires on a certain index.

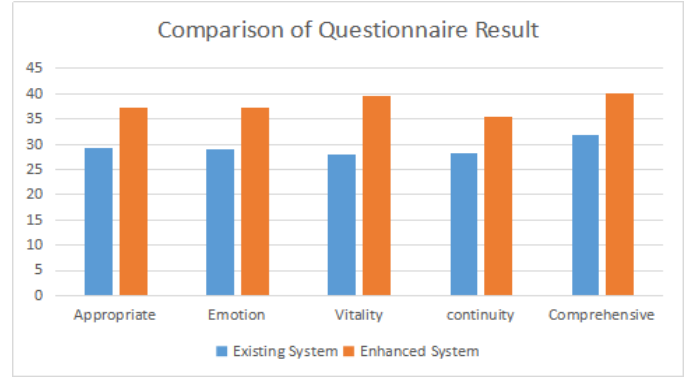


Fig. 19. Comparison of questionnaire results

According to the questionnaire results, the scores of these five indicators of the enhanced system are higher than those of the existing system. Especially the conversion generated by the enhanced system has improved significantly in terms of vitality expression.

VII. CONCLUSION AND FUTURE WORK

To sum up, the objective of this research is aimed at adding emoji stickers to the generated reply sentences for a visual-based emotion expression. Based on Maoyuan's developed conversational system, this research developed a sentimental analysis model and enhanced reply sentences with appropriate emoji based on the sentimental analysis results. An enhanced conversational system with added emoji for reply sentences to the conversation counterpart platform has been developed. The developed prototype system can be run on the WeChat platform.

REFERENCES

- [1] Maoyuan Shen, and Runhe Huang. "A Personal Conversation Assistant Based on Seq2seq with Word2vec Cognitive Map." 2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI). IEEE, 2018.
- [2] Joel Gn. "Emoji as a 'language' of cuteness." First Monday 23.9 (2018).
- [3] Sutskever Ilya, Oriol Vinyals, and Quoc V. Le. "Sequence to sequence learning with neural networks." Advances in neural information processing systems. 2014.
- [4] Rong Xin. "word2vec parameter learning explained." arXiv preprint arXiv:1411.2738 (2014).
- [5] Yan-Yan Zhao, Bing Qin, and Ting Liu. "Sentiment analysis." Journal of Software 21.8 (2010): 1834-1848.
- [6] Sundermeyer Martin, Schlüter Ralf, and Hermann Ney. "LSTM neural networks for language modeling." Thirteenth annual conference of the international speech communication association. 2012.
- [7] Wen Wang, Shufeng Wang, and Honghua Li. "Microblogging sentiment analysis method based on text semantics and expression tendentiousness." Journal of Nanjing University of Science and Technology [ISSN: 1005-9830/CN: 32-1397/N] 6 (2014).
- [8] WeChat Tencent, "2018 WeChat data report", Support.weixin.qq.com, 2019. [Online]. Available: <https://support.weixin.qq.com/cgi-bin/mmsupport-bin/getopendays>. [Accessed: 06- Jun- 2019]