Cross-Lingual Sentiment Research: A comprehensive review of Chinese-English bilingual sentiment analysis

Yuhuang Chen

Data Science and Big Data Technology, XJTLU

10/Apr/2024 @13:00

Abstract

In the task of cross-language sentiment analysis, it is often found that the lack of Chinese corpus makes it difficult to accurately analyze the emotion of Chinese text. To solve this problem, this essay combines machine translation and sentiment analysis technology, and puts forward a comprehensive method to improve the effect of sentiment analysis.

Through the comparison and experimental evaluation of different models, the aggregation model constructed by Seq2Seq and LSTM is selected, and a richer English corpus is used for sentiment analysis to understand the contextual emotion content of the text. The model is trained by using encoder-attention-decoder structure and the improved LSTM structure based on BPTT algorithm. Loss trend, accuracy, precision, recall rate and f1-score were used as performance evaluation metrics. Compared with Naïve Bayes, SVM and logistic regression, our model achieves the highest accuracy of 80.21%, which is superior to these model methods. The method proposed in this essay provides an effective framework for Chinese-English cross-language sentiment analysis. Based on this, we can understand the expression of emotions in different cultural backgrounds, extract emotional features and analyze user comments.

Keywords: Seq2Seq; machine translation; LSTM; sentiment analysis; Chinese-English; cross-language

CONTENTS

List of Acronyms

Term	**Initial Components of the Term**
AI	Artificial Intelligence
ANN	Artificial Neural Network
BPTT	Back Propagation Through Time
CNN	Convolutional Neural Network
FN	False Negative (Observation is positive but is predicted negative.)
FP	False Positive (Observation is negative but is predicted positive.)
GRU	Gated Recurrent Unit
TN	True Negative (Observation is negative and is predicted to be negative.)
TP	True Positive (Observation is positive and is predicted to be positive.)
LSTM	Long Short-Term Memory
RNN	Recurrent Neural Network
SVM	Support Vector Machines
Seq2Seq	Sequence-to-Sequence

Chapter 1 Introduction

1.1 Background

With the third wave of AI technology development, people are moving into the era of artificial intelligence. Nowadays, ChatGPT, BingAI and other chatbots can understand human language and respond to it quickly [1]. Despite the rapid development of artificial intelligence, its emotional understanding in the Chinese filed is still not thorough. English is the Indo-European language family and Chinese is the Sino-Tibetan language family [2]. Therefore, there are significant differences in sentence structure between the two languages. But in fact, most of the mature chatbots on the market today are trained from English corpus, it is very difficult to train a Chinese-based model for analyzing sentiment in Chinese paragraphs. In order to solve this problem, an attempt is made to translate Chinese into English and then conduct sentiment analysis to try a larger corpus.

1.2 Problem Statement

Machine translation and sentiment analysis are important tools in Chinese and English bilingual sentiment analysis. Machine translation expands the scope and applicability of data by translating into different languages. Sentiment analysis can accurately judge the emotion expressed in the text, so as to better grasp the emotional tendency of different language groups [3]. However, most of the models on the market require a great amount of training and time-consuming labor costs, and it is difficult to find a suitable model for cross-language sentiment analysis in small teams. Moreover, users often use informal and abbreviated language to communicate, sentiment analysis of such texts can face the problems of understanding simplification, ambiguous expression, and contextual dependence. At the same time, due to the small amount of Chinese corpus data, it is difficult to conduct sentiment analysis with higher accuracy.

1.3 Purpose of the Coursework

Sentiment analysis using mixed data in Chinese and English has several applications in opinion mining ranging from customer satisfaction to social campaign analysis in multilingual societies. This essay aims to use Seq2Seq model to transform Chinese text into English text, and then use the LSTM model to perform sentiment analysis on the output text. This helps chatbots make use of the rich research results of sentiment analysis in English corpus and improve the accuracy of sentiment analysis of Chinese texts [4]. By translating Chinese into English, more perfect English sentiment analysis methods can be used to promote the research in the field of cross-lingual sentiment analysis.

Chapter 2 Literature Review

2.1 Overview of Current Models

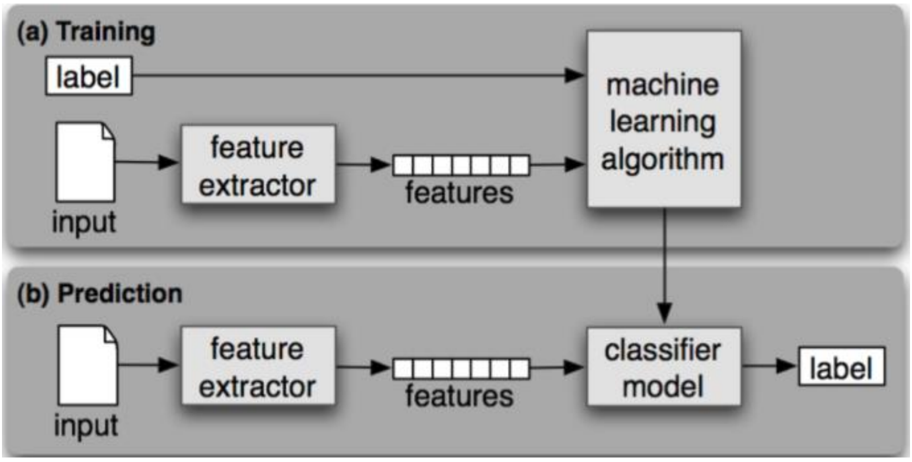
The authors of [5] introduce a multimodal sentiment chat translation dataset, which aims to promote the research of multimodal sentiment analysis and translation. To evaluate multimodal chat translation tasks, they built several benchmark systems based on the Transformer and adapted them with advanced multimodal machine translation and text chat translation models. When investigating bilingual sentiment analysis of texts, consider whether Transformer has enough generalization and usability in machine translation.

Transformer introduces a self-attention mechanism that allows for parallel computing and can efficiently capture long-distance dependencies. Encoder context information is transformed through encoder-decoder structures and used to generate translation results during decoding. This context-aware capability enables the model to better understand and generate output related to previous text [6]. However, the memory and computing power requirements of Transformer increase significantly when processing large datasets, which can make the entire task expensive and time consuming. Because its representation is based on word embeddings, Transformer cannot accurately process words that do not appear in the training data.

A paper published by Google in 2016 illustrates how the translation quality of the Seq2Seq model: approach or exceed all currently published results. Tran et al. [7] proposed using the encoder output of the multilingual Sequence-to-Sequence (Seq2Seq) model obtained by self-supervised training as a language-agnosia representation for retrieving parallel sentence pairs. It is also pointed out that the Seq2Seq can capture the context information better than other models when processing sequence data. Secondly, Seq2Seq model is relatively small, with fewer parameters and computational requirements, which makes it more feasible in some resource-constrained environments. However, it should be noted that when the length of the input sequence exceeds the maximum length set by the model in the training stage, the input sequence needs to be truncated or filled. This can result in loss of information.

Sentiment analysis is a technique for evaluating a sentence or a word based on its emotional content [8]. Machine learning models can automatically learn emotional representations from text by learning from training data that has labeled emotions. It is necessary to review whether machine learning can perform accurate sentiment analysis on this mixed language.

Table 2-1: Machine Learning Process



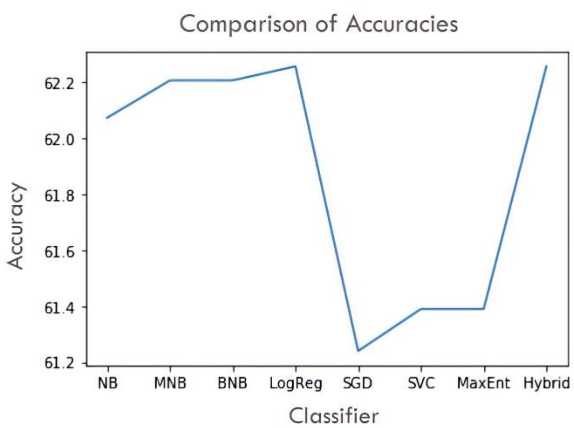
The authors of [9] passed the detailed code test and focus on the accuracy of Naïve Bayes, Logistic Regression and SVM.

i. Naïve Bayes is a classification method based on Bayes's theorem and the assumption of conditional independence of features. It has good performance for high dimensional data and it is not sensitive to lost data. However, the conditional independence of features is assumed to be strict, so it cannot perform well on complex datasets.

ii. Logistic Regression is a statistical learning method used to solve binary classification problems by mapping the output of a linear regression model to a logical function (sigmoid function) to achieve classification

[10]. It has high computational efficiency, strong interpretability, and can handle linearly separable data well. However, the ability to fit the nonlinear feature space is weak and it is susceptible to correlations between features

iii. The goal of Support Vector Machine (SVM) is to find an optimal hyperplane that divides data into different classes by maximizing the margin between classes. SVM can efficiently handle small to medium sized datasets and provide good generalization performance. However, SVM is too sensitive to parameter selection, and the speed of training and prediction on large-scale data is slow.



Classifier Used	Accuracy
Naive Bayes	62.0729
Multinomial Naive Bayes	62.2062
Bernoulli Naive Bayes	62.2062
Logistic Regression	62.2562
SGD	61.2397
SVC Classifier	61.3897
Max Entropy	613897
Hybrid Model	62.2563

Figure 2-1: **Accuracy of Sentiment Analysis based on Machine Learning.**

Note: SVC Classifier = SVM

According to the actual code operation, the accuracy of sentiment analysis based on traditional machine learning is generally **less than 65%**. When a hybrid model of machine learning is tried, there is still no significant improvement in accuracy. Therefore, continue to explore the important branch of Machine Learning - **Deep Learning.**

Deep learning is based on the concept of ANN that learn feature representations of data that learn feature representations of data through multiple layers of processing.

i. In sentiment analysis, Convolutional Neural Network (CNN) can automatically learn the local features and patterns of text, with high computational efficiency and good feature extraction effect for phrases and phrases [11]. However, CNN cannot be as effective as RNN for sentiment analysis tasks that deal with long text or are more context-dependent.

ii. Recurrent Neural Network (RNN) is specifically designed for processing sequence data. The design of the network allows information to be passed from one time step in the network to the next, enabling it to model and learn the timing features and dependencies of the sequence data. RNN can model the order and dependency of texts, so as to better understand the relationship between words in sentences [12]. But traditional RNN have difficulty in processing long sequences of data or texts with long-term dependencies due to the gradient disappearance problem.

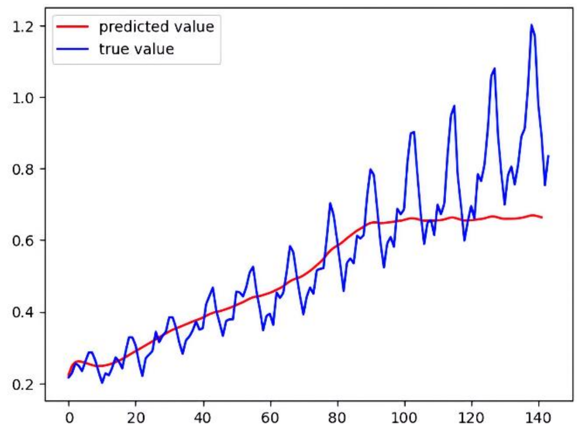


Figure 2-2: **Comparison of True and Predicted Values of Traditional RNN.**

Although RNN is better suited for sentiment analysis than CNN. Long Short-Term Memory (LSTM) improves on the basis of RNN. LSTM is proposed to solve the problem of gradient disappearance and gradient explosion of traditional RNN. LSTM network's core compose input gates, forget gates, output gates. Modulating these gates enable the addition, deletion and retention of information, which enhance the processing of long text sequences. However, the training of LSTM model usually requires a large amount of data and a long training time, and it is difficult to capture global information and is easily affected by local noise.

2.2 Research Gap Identification

In this information-driven era, cross-language sentiment analysis needs to use machine translation to alleviate the difficulty of semantic understanding caused by different languages, and then judge the emotion for sentiment assessment. However, there is still a gap between Chinese and English sentiment analysis technology and the expected direction.

1. Machine translation models lead to the loss of original emotional information. When conducting sentiment analysis after translating Chinese text into English, it is necessary to research how to maintain the accuracy and coherence of information in the process of machine translation [13].
2. Traditional Seq2Seq-based machine translation models are often difficult to capture remote dependencies. It is necessary to consider how to improve the handling of remote dependencies with more efficient encoder-decoder structures.
3. In English and Chinese bilingual sentiment analysis, the acquisition of sentiment labeling data is costly and subjective, and there are many ambiguities in sentiment analysis [14].
4. Sentiment analysis requires not only understanding the polarity of emotion in a single sentence, but also considering the impact of context on emotion. It is necessary to consider how to introduce a better context modeling mechanism into the sentiment analysis model based on LSTM.

Chapter 3 Aims and Objectives

3.1 Aims

The primary aim of this research is to address the complexity of sentiment analysis in both Chinese and English contexts. At the same time, it aims to solve the problem of semantic ambiguity of the two languages.

3.2 Objectives

This essay is specifically carried out in the following aspects:

i. The main objective is to expand the scope of cross-language sentiment research by conducting bilingual sentiment analysis in Chinese and English.

ii. In view of the sparse problem of Chinese corpus, the objective of the research is to use translation technology to reduce the limitations brought by data scarcity.

iii. The last objective is to collect the full format of English abbreviations to replace text content with semantic ambiguity and avoid inaccurate analysis [15].

```
"Whatcha":"What are you",
"luv":"love",
"sux":"sucks",
"shit":"bad",
"tmr":"tomorrow",
"tmrw":"tomorrow",
"u":"you",
"ur":"your",
"k":"okay",
```

***Figure 3-1:* Snapshot of Abbreviations.**

Chapter 4 Methodology

4.1 Research Design

Read the "en-ch" data and preprocess it, then divide the Chinese and English text into words, filter out the sentences with the specified number of words. Then define the encoder and decoder, and add the attention mechanism. After the translation model is defined, the vectorization of emotion comments is carried out, and the sample is trimmed with padding [16]. Configure parameters and train the model. In the end, input sentences for translation, and then translate sentences for sentiment analysis.

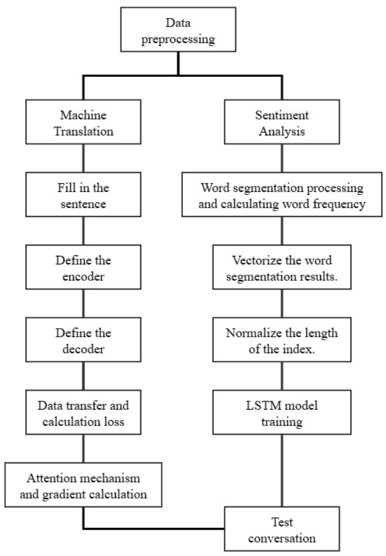


Figure 4-1: **Design Flow Chart.**

4.2 Data Collection

Through the collection of UM-Corpus [17], the most common 20k Chinese-English conversation phrases in daily life are screened out. Combine "dailydialog" [18] and "isear" [19] to create a dataset of positive and negative emotions, consisting of 17k comments respectively. The first $\frac{3}{4}$ of the sentiment dataset is selected as the training set and the rest as the validation set.

The content of People's Daily conversations was randomly intercepted as a test. Select people 1: “嘿，放松点。”, people 2: “我没事。”in this test.

4.3 Algorithms

Machine translation requires mapping an input sequence to an output sequence and taking contextual information into account. Seq2Seq algorithm adopts encoder-attention-decoder structure, which can capture input semantic information and generate output, then use an attention mechanism to store more information. Set $x = \{x_1, x_2, \dots, x_n\}$ is the input, $y = \{y_1, y_2, \dots, y_n\}$ is the output, and y_t is the current output word. The function of Seq2Seq is as follows, with the output dependent on the previous output and input x, which improves understanding of the context.

$$P(y|x) = \prod_{t=1}^{m} logP(y_1|y_1, y_2, \dots, y_{t-1}, x)$$

Sentiment analysis tasks need to consider long-term dependencies between words in a text. The BPTT algorithm calculates the error gradients of each time step through the backpropagation algorithm, and then uses these gradients to update the network parameters [20]. On this basis, LSTM algorithm overcomes the gradient explosion problem by gradient clipping technology [21].

4.4 Model Training and Validation

In machine translation, the operation during forward propagation is recorded, then loss value is calculated using the loss function, and backpropagation is performed to update the parameters. Finally, select common and rare conversations in daily life for validation to ensure stability. Then in sentiment analysis, through compile, the evaluation index of the constructed model is set to “accuracy” [22]. The model is used to predict the data of the validation set, obtain the predicted probability value, then convert it into the predicted category, and select the category with the highest probability as the prediction result.

4.5 Performance Evaluation Metrics

By drawing the change of the loss value with the number of training iterations, it can be observed whether the loss value gradually converges during the training process of the machine translation model.

In sentiment analysis, the following metrics are defined (terms are represented in List of Acronyms):

A. Accuracy

Accuracy is a ratio of correctly predicted observations to the total observations.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

B. Precision

Precision is the fraction of relevant instances among the retrieved instances. It is basically used as the measure of relevance [23].

$$Precision = \frac{TP}{TP + FP}$$

C. Recall

Recall is the fraction of relevant instances that have been retrieved over the total number of instances.

$$Recall = \frac{TP}{TP + FN}$$

D. F1-Score

It is the weighted average of precision and recall [24].

$$F1 - Score = \frac{2 * Recall * Precision}{Recall + Precision}$$

Chapter 5 Outcomes and Deliverables

5.1 Model Performance

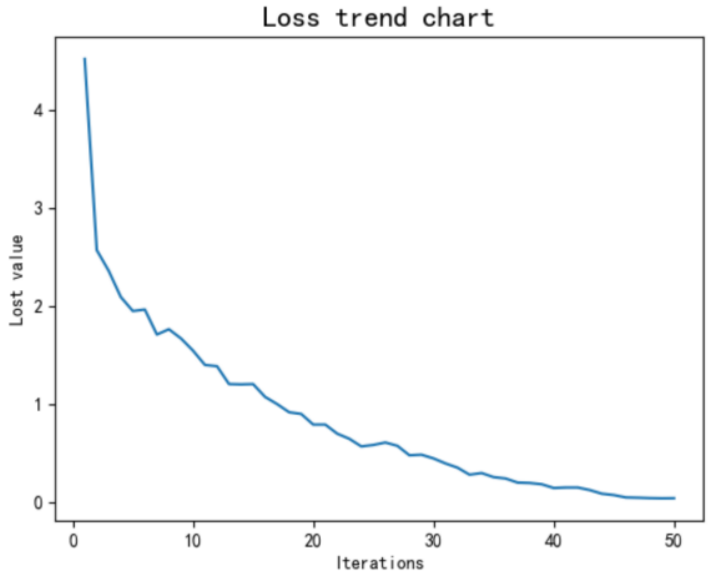


Figure 5-1: **Loss Value Diagram.**

As the number of iterations increases, the loss value of the model on the machine translation training set gradually decreases and becomes stable, which means that the prediction results are more accurate.

```
The accuracy of the test set is:  0.8020718536477849
Accuracy rate, recall rate and F1 value are respectively:
              precision    recall  f1-score   support

           0       0.83      0.80      0.82      4974
           1       0.77      0.81      0.79      4100

    accuracy                           0.80      9074
   macro avg       0.80      0.80      0.80      9074
weighted avg       0.80      0.80      0.80      9074

The confusion matrix is:
[[3975  999]
 [ 797 3303]]
```

Figure 5-2: **Sentiment Analysis Performance.**

Figure 5-2 illustrates the accuracy of the model, which exceeds 80% on various metrics and shows very reliable performance in both categories.

5.2 Practical Implications

The built model can be deployed on social media platforms, and by automatically translating comments and performing sentiment analysis, it will change the way user feedback is handled. For multinational companies, understanding consumer sentiment in different countries is crucial for marketing, helping companies to better analyze consumer sentiment in different linguistic contexts.

5.3 Limitations and Challenges

Sentiment analysis needs to consider the contextual information of the text, and there are semantic gaps between different languages. However, in machine translation, the multi-semantic information translated may affect the understanding of emotions. When the trained bilingual sentiment analysis model is applied to data in different fields, it may encounter the problem of poor generalization ability [25].

Chapter 6 Conclusions

6.1 Data Analysis

Initially, special characters are removed from the machine translation dataset and word pairs [English, Chinese] are returned. Create a word index (word→id) and a reverse word index. After processing the outliers in the data, fill each sentence to its maximum length.

```
The output before preprocessing is:
Input language: Word mapping index
1 ----> <start>
89 ----> 你的狗在哪里？
2 ----> <end>
Target language: Word mapping index
1 ----> <start>
67 ----> where
9 ----> s
27 ----> your
87 ----> dog
5 ----> ?
2 ----> <end>
```

Figure 6-1: **Word Mapping Index.**

Define the input of the encoder and LSTM layer to obtain the output and status information. Then define the input of the decoder, take the state of the encoder output as the initial state of the initial decoder, add the full connection layer and define the model. Data is input to the encoder to obtain the output and hidden state. Pass the information to the decoder, which returns the predicted and hidden state and passes the hidden state back to the model to calculate the loss value. The attention mechanism determines the next input of the decoder, calculates the gradient and applies it to the optimizer and backpropagation. Then read the corpus of positive and negative emotions, continue word segmentation processing on the review data, calculate the frequency, vectorize the results of word segmentation, and standardize the length of the index. Set the parameters of LSTM model and train it. Then enter the following conversation to test.

```
1/1 [==============================] - 0s 38ms/step
1/1 [==============================] - 0s 31ms/step
Text 1:                   嘿，放松点。
Translated Text 1:        hey , relax .
Emotion:                  Positive

Text 2:                   我沒事。
Translated Text 2:        i am okay .
Emotion:                  Positive
```

6.2 *Interpretation of Results*

Emotion is universal in language, but there are differences in the expression of emotion in languages. Through bilingual sentiment analysis, we can understand the ways of emotion transmission between different languages, which is helpful to understand the ways of emotion expression in different cultures. Bilingual sentiment analysis can also assess the quality of translation. If the emotional disposition remains consistent after translation, the translation can be considered accurate. This is useful for the task of extracting emotional features from multilingual data and analyzing user comments.

6.3 *Comparison with Existing Methods*

In this section, Naïve Bayes, SVM and Logistic Regression were selected as baseline methods for comparison experiments. To avoid limitations due to insufficient of "en-ch" corpus, use Google Translate and then apply the above methods to sentiment analysis [26]. Precision, Recall and F1-score were chosen to comprehensively evaluate the reliability of the model. The authors of [9] have tested previous methods and obtained accurate probabilities. Compare them based on the results.

A. Naïve Bayes

```
Original Naive Bayes

              precision    recall  f1-score   support

    positive       0.59      0.78      0.67      2979
    negative       0.68      0.46      0.55      3022

    accuracy                           0.62      6001
   macro avg       0.64      0.62      0.61      6001
weighted avg       0.64      0.62      0.61      6001
```

B. SVM

```
Support vector classifier

              precision    recall  f1-score   support

    positive       0.58      0.79      0.67      2979
    negative       0.68      0.44      0.53      3022

    accuracy                           0.61      6001
   macro avg       0.63      0.62      0.60      6001
weighted avg       0.63      0.61      0.60      6001
```

C. Logistic Regression

```
Logistic regression

              precision    recall  f1-score   support

    positive       0.59      0.78      0.67      2979
    negative       0.68      0.47      0.55      3022

    accuracy                           0.62      6001
   macro avg       0.64      0.62      0.61      6001
weighted avg       0.64      0.62      0.61      6001
```

As can be seen from the comparison between ***Figure 5-2*** and the above results, the support degree of our method is much higher than that of the comparison methods, which indicates that our method has been tested in a wider range of examples, and has demonstrated high robustness in various performances.

The accuracies of all the compared models and OURS together are as follows:

Table 6-1: Classifier Accuracies.

Classifier	***Accuracy***
Naïve Bayes	62.0729
SVM	61.3897
Logistic Regression	62.2562
OURS	**80.2072**

6.4 Conclusions

Some of the key features presented in the field of cross-language sentiment analysis help us gain a broader understanding of the subject. The research conducted is unique in several ways compared to similar work. The methodology implements an aggregated model of translation and sentiment analysis based on the Seq2Seq and LSTM models and provides the highest accuracy of 80.21% in the comparison of different methods. If the word is in Chinese, machine translation will translate it directly into English to avoid the lack of a Chinese sentiment analysis corpus.

6.5 Future Works

Our research provides a framework for potential avenues of exploration. If we look to the future to strengthen the research, we can design more objective emotion labeling techniques that can produce more reliable sentiment analysis results. Meanwhile, in-depth research on improving strategies for translation information loss to mitigate its impact on sentiment analysis results can enhance the robustness of our approach. In addition, this research can be extended to multiple regional languages and can be transformed into a multilingual sentiment analysis problem to make it more versatile.

REFERENCE

[1] V. Scotti, L. Sbattella, and R. Tedesco, “A Primer on Seq2Seq Models for Generative Chatbots,” *ACM Computing Surveys*, vol. 56, no. 3, pp. 1–58, Oct. 2023, doi: https://doi.org/10.1145/3604281 .

[2] L. Ling and T. Sepora, “The Differences between English and Chinese Language Sentence Structure and their Impacts to English-Chinese Machine Translation,” 2013.

[3] O. Amezian, S. Hajbi, Z. S. Houssaini and Y. Chihab, "Training an LSTM-based Seq2Seq Model on a Moroccan Biscript Lexicon," 2023 9th International Conference on Optimization and Applications (ICOA), AbuDhabi, United Arab Emirates, 2023, pp. 1-6, doi: 10.1109/ICOA58279.2023.10308821.

[4] L. Wang and L. Wang, “A Case Study of Chinese Sentiment Analysis of Social Media Reviews Based on LSTM,” *SHS Web of Conferences*, vol. 157, p. 04012, 2023, doi: https://doi.org/10.1051/shsconf/202315704012 .

[5] Y. Liang, F. Meng, J. Xu, Y. Chen, and J. Zhou, “MSCTD: A Multimodal Sentiment Chat Translation Dataset,” 2022.

[6] S. Takase and S. Kiyono, “Lessons on Parameter Sharing across Layers in Transformers,” Apr. 2021.

[7] C. Tran, F. Ai, Y. Tang, X. Li, and J. Gu, “Cross-lingual Retrieval for Iterative Self-Supervised Training,” 2020.

[8] V. Yadav, P. Verma, and V. Katiyar, “Long short term memory (LSTM) model for sentiment analysis in social data for e-commerce products reviews in Hindi languages,” International Journal of Information Technology, vol. 15, no. 2, pp. 759–772, Jun. 2022, doi: https://doi.org/10.1007/s41870-022-01010-y .

[9] S. Goel, S. Rahman, and A. Joshi, “Sentiment Analysis of Hindi-English Code-mixed Social Media Text.” 2021.

[10] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, “Language Models are Unsupervised Multitask Learners.” 2019.

[11] H. Nankani, H. Dutta, H. Shrivastava, P. V. N. S. Rama Krishna, D. Mahata, and R. R. Shah, “Multilingual Sentiment Analysis,” *Algorithms for Intelligent Systems*, pp. 193–236, 2020, doi: https://doi.org/10.1007/978-981-15-1216-2_8.

[12] A. Joshi, A. Prabhu, M. Shrivastava, and V. Varma, “Towards Sub-Word Level Compositions for Sentiment Analysis of Hindi-English Code Mixed Text,” 2016.

[13] S. J. Kalyanshetti, M. S. Jagtap, A. U. Kale and P. Waghmare, "Comparative Study of Different Models for Language Translation," 2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA, Pune, India, 2022, pp. 1-4, doi: 10.1109/ICCUBEA54992.2022.10011127.

[14] J. Barnes, R. Klinger, and S. Schulte, “Bilingual Sentiment Embeddings: Joint Projection of Sentiment Across Languages,” Association for Computational Linguistics, 2018.

[15] S. J. Kalyanshetti, M. S. Jagtap, A. U. Kale, and P. Waghmare, “Comparative Study of Different Models for Language Translation,” 2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA, Aug. 2022, Published, doi:10.1109/iccubea54992.2022.10011127.

[16] A. Baliyan, A. Batra, and S. Singh, “Multilingual Sentiment Analysis using RNN-LSTM and Neural Machine Translation,” 2021.

[17] Liang Tian, Derek F. Wong, Lidia S. Chao, Paulo Quaresma, Francisco Oliveira, Shuo Li, Yiming Wang, Yi Lu, "UM-Corpus: A Large English-Chinese Parallel Corpus for Statistical Machine Translation". Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, Iceland, 2014.

[18] “Yanran’s Attic,” Yanran’s Attic. http://yanran.li/dailydialog.html (accessed Apr. 21, 2024).

[19] Affective-sciences.org, 2024. http://www.affective-sciences.org/index.php/download_file/view/395/296/ .

[20] "Calculus on Computational Graphs: Backpropagation -- colah's blog," colah.github.io. https://colah.github.io/posts/2015-08-Backprop/

[21] M. A. Hasan, "Ensemble Language Models for Multilingual Sentiment Analysis," arXiv.org, Mar. 09, 2024. https://arxiv.org/abs/2403.06060 (accessed Apr. 22, 2024).

[22] J. Kocoń, "Deep Emotions Across Languages: A Novel Approach for Sentiment Propagation in Multilingual WordNets," 2023 IEEE International Conference on Data Mining Workshops (ICDMW), Dec. 2023, doi: https://doi.org/10.1109/icdmw60847.2023.00101 .

[23] M. Waseem, "Classification In Machine Learning | Classification Algorithms," Edureka, Dec. 04, 2019. https://www.edureka.co/blog/classification-in-machine-learning/ .

[24] Geeksforgeeks, "Confusion Matrix in Machine Learning - GeeksforGeeks," GeeksforGeeks, Feb. 07, 2018. https://www.geeksforgeeks.org/confusion-matrix-machine-learning/ .

[25] A. Rath, B. Hridaya, D. Vimala, and J. George, "Multilingual Sentiment Analysis of YouTube Live Stream using Machine Translation and Transformer in NLP," 2022 International Conference on Trends in Quantum Computing and Emerging Business Technologies (TQCEBT), Oct. 2022, doi: https://doi.org/10.1109/tqcebt54229.2022.10041483 .

[26] B. H.R.A.M.K, A. M.M.Z, A. Gamage, D. S. S.H.D.Y, A. M.A.R, and A. Caldera, "SentiNet: A Robust and Multilingual Sentiment Analysis System with Transfer Learning and Adversarial Training Techniques," 2023 5th International Conference on Advancements in Computing (ICAC), Dec. 2023, doi: https://doi.org/10.1109/icac60630.2023.10417466 .