

Review of Text Style Transfer Based on Deep Learning

Xiangyang Li¹,Guo Pu¹,Keyu Ming¹,Pu Li¹,Jie Wang¹,YuXuan Wang¹

Peking University

{xianyangli,guopu,kae_ming,spurslipu,jenny.wang,Lovedn4ever}@pku.edu.cn

Abstract

Text style transfer is a hot issue in recent natural language processing, which mainly studies the text to adapt to different specific situations, audiences and purposes by making some changes. The style of the text usually includes many aspects such as morphology, grammar, emotion, complexity, fluency, tense, tone and so on. In the traditional text style transfer model, the text style is generally relied on by experts knowledge and hand-designed rules, but with the application of deep learning in the field of natural language processing, the text style transfer method based on deep learning Started to be heavily researched. In recent years, text style transfer is becoming a hot issue in natural language processing research. This article summarizes the research on the text style transfer model based on deep learning in recent years, and summarizes, analyzes and compares the main research directions and progress. In addition, the article also introduces public data sets and evaluation indicators commonly used for text style transfer. Finally, the existing characteristics of the text style transfer model are summarized, and the future development trend of the text style transfer model based on deep learning is analyzed and forecasted.

1 Introduction

Since 2006, deep learning has gradually become a hot research direction of machine learning, and has been widely concerned by industry and academia. Deep learning has made breakthroughs in computer vision, image processing, and natural language processing. For example, style transfer is a hot topic that has been studied in the field of artificial intelligence. Its research results are mainly distributed in the field of computer vision and natural language processing. In recent years, style transfer in the field of computer vision has been fully studied. In 2015,(Gatys et al., 2015)pioneered a transfer of image styles based on convolutional neural networks, and found that convolutional neural networks can be used to represent the "content abstract feature representation" and "style abstract feature representation" After separation, the separated "style abstract feature representation" is applied to the new content, and images with very different styles can be obtained. "Style abstract feature representation" itself includes many different elements, including "texture", "color", "line", "shape" and so on. After this, a large number of studies on the transfer of image styles have emerged, and have led to the research of generation techniques in the field of computer vision, such as generating adversarial networks (Goodfellow et al., 2014) and automatic variational encoders(Kingma et al., 2013).

At present, deep learning has made gratifying progress in the task of image style transfer in the field of computer vision. In the direction of natural language processing, deep learning technology has played a significant advantage in natural language processing tasks such as entity relationship extraction(Miwa M et al., 2016), event extraction(Lin CY et al., 2016), machine translation(Cho K et al., 2014), and automatic question answering(Zhang P et al., 2018). However, as a new field in natural language processing, the task of text style transfer still faces great challenges. Compared with the

image style transfer, the "style abstract feature representation" of the text is more elusive, and it is difficult to strip it out by explicit methods. Therefore, the current text style transfer is still in the exploration stage. However, at many international conferences this year, the text style model combined with deep learning has emerged, and it has caused a new research boom at related international top conferences (ACAL, IJCAI, AAI, SIGIR, KDD).

At present, the predecessors already have some overviews on the transfer of image styles(Y.Jing et al., 2017). However, so far, almost no one has compiled a summary of the achievements of text style transfer. Therefore, in order to systematically review the research results of the text style transfer model based on deep learning technology, we collated and analyzed the top journals in the field of natural language processing, artificial intelligence, machine learning and data mining in the past five years in the field of text style. The research results summarize and summarize the existing related methods, and compare and discuss the advantages and disadvantages of each method.

The article is organized as follows: Section II of this article introduces the transfer of text style, and gives the basic framework of the formal definition of text style transfer and the text style transfer model. Sections 3 and 5 summarize various methods of combining deep learning text style transfer models. Section 6 compares and summarizes the public datasets used by the academic community in the transfer of text styles. Section 7 introduces the evaluation indicators that introduce text style transfer. Section VIII makes a preliminary exploration of the research directions that deserve attention in the future. Section IX summarizes the full text.

2 Introduction to Text Style Transfer

In recent years, the natural language generation task (NLG) has received more and more attention from researchers, including tasks such as dialogue generation(Reina et al., 2017) , machine translation (Dzmitry Bahdanau et al., 2014), and automatic summarization (Sumit Chopra 2016). The purpose of text style transfer is to give The defined sentences are converted according to the required style, such as the emotion from positive to negative, the writing style is switched from the spoken style to the Shakespeare style, and the tone style is transferred from formal to informal. While transferring these styles, and keep the text content unchanged. In terms of application, the text style also shows great value, such as text normative checking, robot chat dialogue, anonymous writing, and paper review.

In the development speed of the research field, the field of text style transfer should lag far behind the transfer of image style. There are three main reasons for this. One is the lack of a large number of parallel corpora, which has led most of the research mainly to unsupervised learning methods. Due to the lack of supervised information, the resulting information has successfully achieved a change between styles, but the content meaning Changes have taken place, which is a major problem arising from the early text style transfer. Another problem is the discretized representation of text, because the discretized encoded representation of the text causes the text data to be unguided when it is derived, and in the image field, it is generally only necessary to go back through the gradient. Good implementation of gradient descent. The third reason is the uncertainty of the evaluation indicators. Generally, for binary styles such as "emotion", you can identify them by training a good binary classifier, but it is difficult to judge the "writing style". Style, it is difficult to discriminate by training a semantic classifier. At this time, it can only be judged by the "semantic fluency", "BLEU" value, and manual identification. This will lead to confusion in the evaluation criteria and cannot judge the model Good or bad.

In recent years, the use of neural networks for text style conversion has been extensively studied. A common method is to use VAE (Automatic Variation Encoder) to separate content and style in the latent space, and then adjust the separated style. The adjustment method is generally through the style classifier or the generation of confrontation network. And then pass the sentence through the decoder to generate another style of sentence. Literature (Hu et al., 2017; Fu et al., 2018;Shen et al., 2017; Yang et al., 2018; Gong et al., 2019; Lin et al., 2017)has conducted in-depth research in this direction. (Li et al., 2018;Xu et al., 2018; Zhang et al., 2018c)by filtering the input of sentences, In order to

achieve the separation of text content and style. (Prabhumoye et al., 2018) proposes to perform text style transfer through the reverse translation method by using a denoising automatic variational encoder. In addition, due to the successful application of sequence model (Sutskever et al., 2014) in the field of natural language processing, most scholars began to use sequence model to study text style transfer.

Based on the different learning methods, this article divides the text style transfer model methods into three categories, namely supervised learning, semi-supervised learning, and unsupervised learning. Supervised learning mainly uses parallel data sets for style transfer, which is characterized by high accuracy and good style transfer effect. The data source of the unsupervised learning method is mixed with parallel data and non-parallel data. Generally, the style representation is learned from the style conversion in the parallel data, and then it is used in the text style transfer of the non-parallel data to solve the non-parallel data. Text style migration issue. The unsupervised learning method only has non-parallel data, and text generation is based entirely on non-parallel data sets. It can show good results in many potential fields where there is no parallel text, but the learning difficulty is also the most difficult.

3 Supervised model

The task of text style transfer in supervised learning usually refers to the method of using parallel corpora in model training.

(Jhamtani et al., 2017) is an important study of the earlier text style transfer task, which proposes a method for automatically transferring text from modern English to Shakespeare English, which uses an end-to-end trainable with pointers that allow copy operations Neural network model.

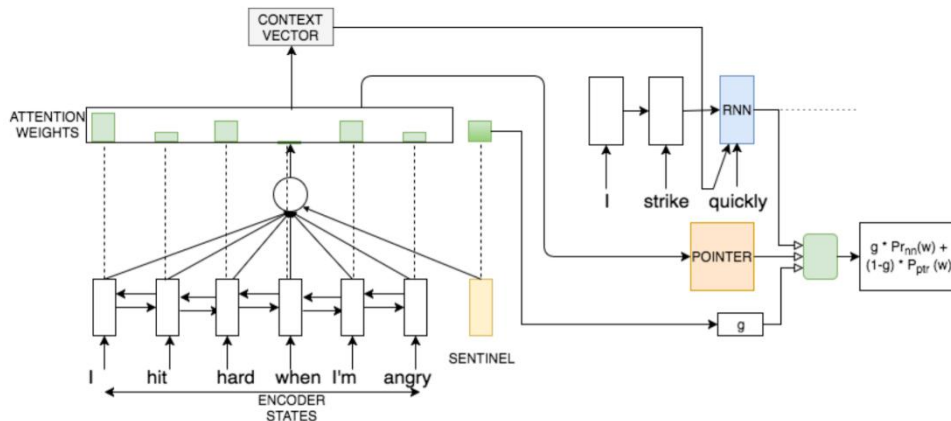


Figure 1: Attention-based sequence-to-sequence model combined with pointer network

The model (as shown in the above figure) combines attention-based seq2seq and pointer network (CopyNet). The encoder part uses bi-LSTM, but unlike the conventional concatenate two-direction vector, it directly performs the add operation, which reduces the parameters. the amount. In order to reduce the emergence of UNK, the study used CopyNet. The model uses a modern Shakespeare dictionary to provide candidate words for the pointer network, which belongs to a parallel corpus.

The data set used in this study is a line-by-line modern interpretation of 16 of 36 Shakespeare 's plays, of which the training set contains 18,395 sentences in 14 plays, and the last play, "Romeo and Juliet," which contains 1,462 sentences, is the test set. . In order to solve the problem of insufficient parallel data, the author combined with additional text to pre-train an external dictionary representation (embeddings) from Shakespeare English to modern English and explained that the shared dictionary table on the source language side and the target language side is beneficial to improve performance. After experiments, the method achieved a BLEU score of 31+, which is about 6 points higher than the strongest benchmark MOSES.

(Carlson K et al., 2018) pointed out that the style of an article that can be perceived is composed of many characteristics, including the length of the sentence, the active or passive voice, the level of vocabulary, the degree of tone, and the formality of the language. However, most text styles at present Migration research does not explicitly consider any of the above aspects, but only a broader style.

To do a more detailed study of style, these systems need parallel data to train and test the results, and the current parallel corpus for style is in a state of emergency. Therefore, the study collected a large, undeveloped parallel text data set-the Bible translation corpus. The study shows that its data set is conducive to the model's generalization of style learning, because each version of the Bible reflects A unique style. Each version exceeds 31,000 sessions and can generate more than 1.5 million unique parallel training data.

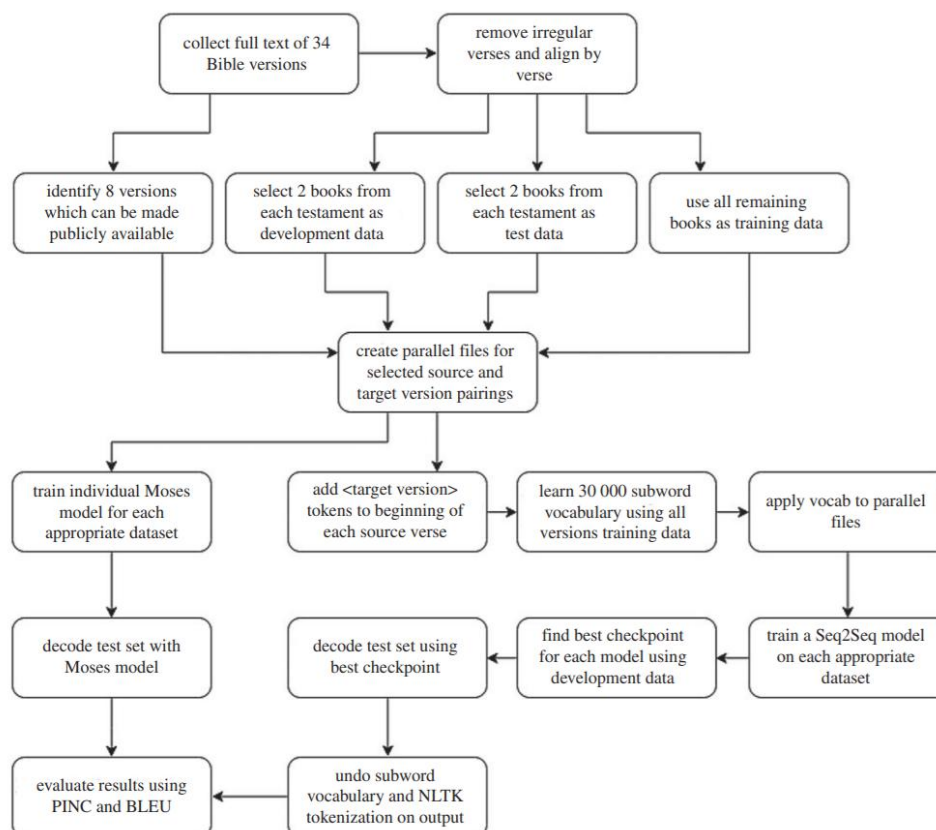


Figure2: The complete flow of the Bible data set from construction to experiment

The study used two models, MOSES and Seq2Seq, of which MOSES is an open source data machine translation toolkit, and Seq2Seq uses a multi-layer cyclic neural network encoder and a multiple-cycle neural network decoder with attention mechanism. In the evaluation, the study used two indicators, BLEU and PINC, to make a comprehensive evaluation. After experiments, both MOSE and Seq2Seq migrated the text to the target style. Overall, MOES achieved a higher BLEU score. When the task needed to make fewer changes to the original sentence to meet the target, the advantages of this BLEU would be more Big. When the task requires more changes to the original sentence, Seq2Seq is better than MOSES in the five most important tests.

(Harnessing et al., 2019) mainly focused on the study of the formality of text (Formality), and proposed a method to transfer the style from informal text to formal text. The study shows that when the current parallel corpus is very small, using a large neural network model that has been pre-trained on a large-scale corpus and has learned general language knowledge will be effective, and the introduction of rules effectively reduces the complexity of the data. The model makes it easier for the model to learn complex patterns. This study shows how to effectively use a pre-trained network with simple rules (GPT-2 model in this study) for formal style transfer tasks.

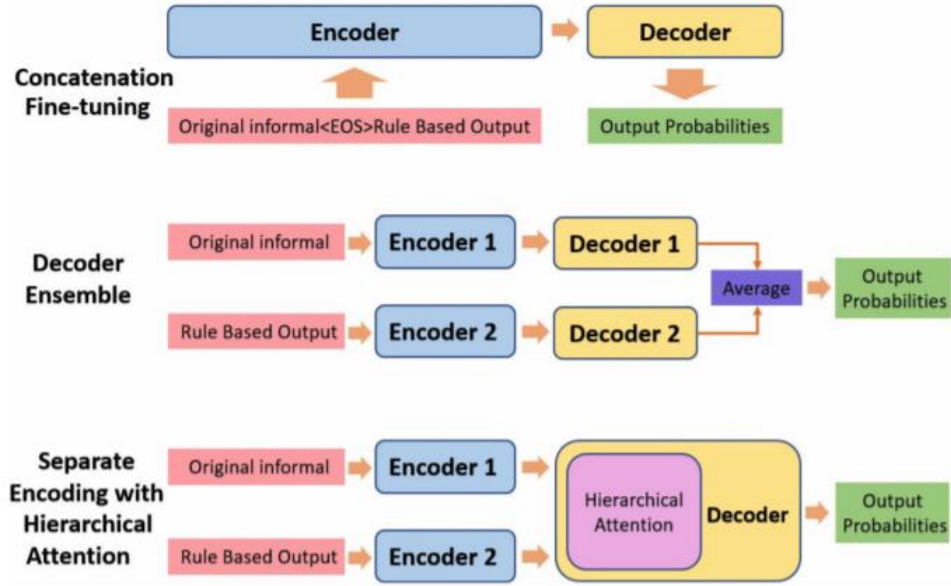


Figure3: Three different ways to incorporate rules into pre-trained models

This study analyzes three methods of using the rules in the GPT-2 model (as shown in the above figure): 1) the initial informal sentence and its pre-processed form are spliced and sent to the encoder; 2) in the inference stage Integrate the two models: one takes the original informal text as input, and the other takes the preprocessed text as input; 3) uses two encoders to encode the initial informal text and the rule preprocessed text, and then uses The vocabulary level and sentence level level attention mechanism to aggregate information.

4 Semi-supervised model

Semi-supervised learning is a recently proposed text transfer model. The semi-supervised learning method combines small-scale parallel data sets (parallel data) and large-scale non-parallel data sets (nonparallel data) to solve the problem of insufficient parallel data sets.

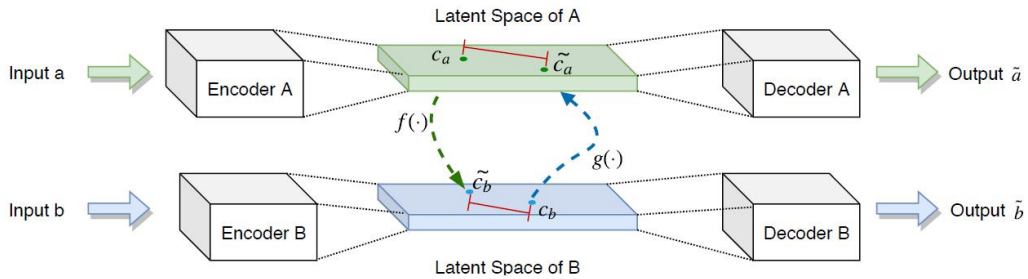


Figure 4: Real-term cross-prediction between two styles of DVE models in semi-supervised text style transfer

(Shang et al., 2019) pointed out that the method based on the standard S2S (sequence-to-sequence) proposed latent space cross prediction method (Cross Projection in Latent Space) realized the function of style conversion between different style data sets. As shown in the figure above, data sets A and B represent two non-parallel style data sets, which are coded and decoded on the two data sets using a denoising automatic encoder (DAE). Set the cross-prediction function on the latent space, and enhance the performance of the cross-mapping function by setting the non-parallel data set and the parallel data set loss function loop. Finally, by inputting from the Encoder module of Style A and outputting from the Decoder module of Style B through the cross prediction function, the text style transfer is realized.

5 Unsupervised model

Unsupervised learning does not use paired source data and target data, and generates text based on non-parallel data sets. It can show good results in many potential fields where parallel text does not exist.

(Keith Carlson et al., 2018) based on the inspiration of Zero-Shot Translation, Zero-Shot Style Transfer is proposed. In this paper, the zero-sample style transfer is converted into a single machine translation problem, and Based on this, a recurrent neural network (RNN) model based on S2S (sequence-to-sequence) architecture is created, as shown in the following figure.

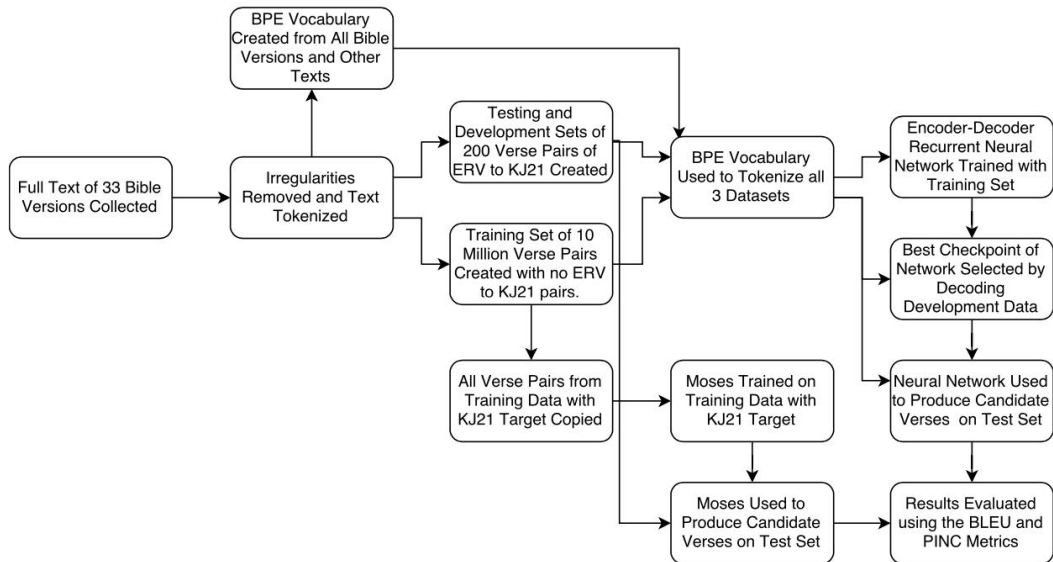


Figure 4:Zero shot style network structure

The model uses a multi-layer recurrent neural network (RNN) encoder and decoder. The encoder and decoder are three-layer bidirectional recurrent neural networks. Each layer has 512 layers of LSTM units, and each layer uses residuals. even. As shown in the figure, the training process of the model is that 64 pairs of poems are randomly selected from the training set. Each source sentence and target sentence are extracted as 100 symbols. These symbols are fed into the encoder. Once the entire sentence is put into the encoder The decoder will be put into a special symbol signal, telling it that it should start generating paraphrases. During training, except for the first time step, other time step decoders input the correct a priori words in the target sentence. The parameters of the model are updated using Adam. The checkpoint is stored every 5000 steps. This checkpoint is used to decode the development set This checkpoint with the best average number of sliding windows and 3 average BLEU scores is used to infer the test set, so as to realize the style transfer of the text.

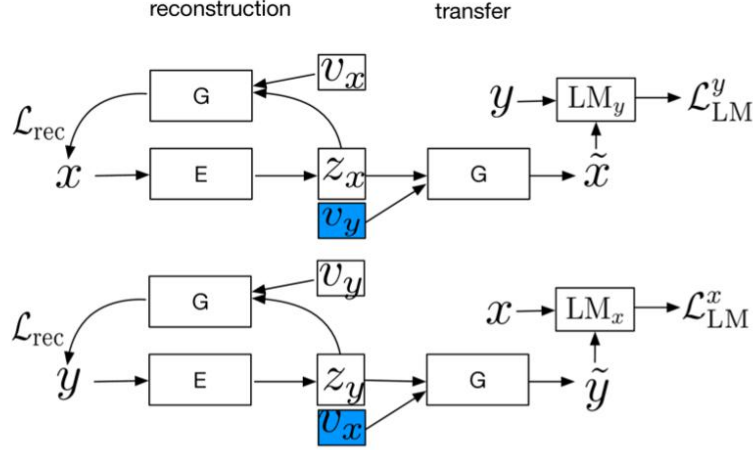


Figure 5:GAN style network structure

In addition to adopting the above-mentioned unsupervised learning based on S2S mechanism, GAN (Generative Adversarial Networks) is also commonly used in style transfer systems. This model usually uses a binary discriminator to ensure that the translated sentence is similar to the sentence in the target domain, but there is also a difficulty. The loss information provided by the discriminator may be unstable, sometimes Not enough to train the generator to produce fluent language. Therefore, the literature (Yang et al., 2018) pointed out that using the language model of the target domain as a discriminator provides a richer and more stable symbol level during the learning process Feedback, training the generator to minimize the NLL (negative log likelihood) of the sentences generated by the language model evaluation. By using continuous estimates of discrete sampling under the generator, we can use back propagation to train our model in an end-to-end manner In addition, the experimental results in this article show that when the language model is used as a structured classifier, the adversarial steps can be abandoned during the training process, thereby making the process more stable. The model structure is shown above.

6 Common dataset

Because the purpose of style transfer is to transfer the text style of some sentences to another large corpus and apply it, the commonly used data sets for style transfer are mainly public data sets. Self-built data sets are not only time-consuming and labor-intensive, but also because of migration effects and practical application considerations, public data sets are generally a better choice. However, for specific applications or research purposes, many scholars also choose to use self-built data sets for research.

Dataset	Language	Size	Access
Wikipedia Corpus	English/Chinese	1,900,000,000 words	free
Yelp review	English	5,200,000 review	free
SST	English	239,232 words	free
IMDB	English	50,000 review	free
Amazon	English	35,000,000 review	free
Twitter	English	11109 text	free

Table 1: Common dataset.

6.1 Wikipedia Corpus

(Zhu et al., 2013) The data set is a collection of Wikipedia full text. It contains nearly 1.9 billion words from more than 4 million articles. This data set allows you to search through a part of a word, phrase, or paragraph itself to get a corpus related to it.

Wikipedia Corpus	Contents
Articles	4,400,000
Words	1,900,000,000
Size	20MB

Table 2: Wikipedia Corpus.

6.3 Yelp Review

Yelp is a famous American merchant review website. Founded in 2004, it includes merchants from restaurants, shopping centers, hotels, and tourism in various areas. Users can rate merchants on the Yelp website, submit comments, and exchange shopping experiences.

Yelp published an open dataset on Github for learning purposes-Yelp reviews(Shen et al., 2017). It contains 5,200,000 comments, 174,000 commercial attributes, and 200,000 photos. The latest updated data is 2017(Li et al., 2017).

6.4 Stanford Sentiment Treebank(SST)

(Socher et al., 2013)The Stanford Sentiment Treebank (SST) dataset is published by the Stanford University ' s NLP group and contains a total of 239,232 sentences and phrases. It also resumes a deep learning model based on the complete representation of sentence structure, which can judge the sentiment of sentences according to the composition of words, which is different from most sentiment judgment systems that ignore the order of words.

SST was released in 2013 and is widely used in sentiment classification problems, and each of the sentence analysis nodes has fine-grained sentiment annotations.

6.5 Other

When studying vocabulary-level tags, we can also directly use a vocabulary in a specific field to train vocabulary, treat it as a sentence and evaluate it. For example, a dictionary containing 2700 words with emotion tags released(Wilson et al., 2005).A feasible application direction of style transfer is to transfer the style of an article to the style of a specific author, so the author's articles form a corpus.

- Bible(Carlson et al., 2015)
- Shakespeare Collection(Jhamtani et al., 2017)
- Ancient Poetry and Literature Website(Ming et al., 2019)

6.6 Summary

It is relatively easy to obtain corpus data sets. Various public data sets, various large-scale websites, various corpus websites, etc. can all become corpus data sources. The use of data sets should vary with the purpose of the research and the direction of application. When testing the model, you can test the model on a variety of large public data sets and observe its generalization ability, such as (YaoFu et al., 2019) At the same time, you can also choose a suitable data set for a certain topic, or even try to provide a self-built data set(Shang et al., 2019).

7 Evaluation indicators and evaluation methods

During the training process, effective evaluation indicators can provide a valuable reference to the output of the model, helping researchers better improve the model based on the feedback content. This section first summarizes the research related to text style transfer evaluation, then summarizes the commonly used evaluation indicators and calculation methods used in text style transfer, and finally discusses the difficulties and problems of text transfer evaluation.

7.1 Research status of text style transfer assessment

At present, compared with the study of the rich text style transfer model, the research on the evaluation method for this task is still lacking. (Yamshchikov et al., 2019) conducted a study on the semantic retention (Similarity) of the pre-migration text for the text after style transfer, and compared the POS-Distance, Word overlap, Cosine similarity (GolVe, FastText)), WMD, BLEU, ROUGE and other commonly used distance measurement methods in the field of NLP are consistent with manual evaluation scores, which proves that in these methods, no single indicator can reflect human evaluation well. In addition, it is pointed out that by combining POS-distance, WMD and BLEU can better measure the semantic retention of the migrated text. (Remi Mir et al.,2019)based on the Yelp user food review data set, studied the text style transfer evaluation method. The study divided the evaluation indicators into style transfer intensity (STI, Style Transfer Itensity) and content retention (CP, Content Preservation) And three aspects of NL, Natualness, different automatic calculation methods are designed from all aspects.

The experimental results were verified on three classic text style transfer models CAAE(Shen et al.,2017), ARAE(Zhao et al.,2018), DAR(Li et al., 2018), and found that the consistency of EMD and manual scoring is the highest for style transfer intensity; for content retention, In the case of style masking the text, the WMD calculation score is the most reliable; for naturalness, the perplexity commonly used by previous people has no correlation with the artificial scoring results, and the adversarial classifier works better. (Pang et al., 2019) discussed the difficulties of automatic evaluation of text style transfer in the real world. The study discussed the three aspects of the task difficulty of the text itself, the difficulty of the corpus used for transfer and the difficulty of score calculation, and proposed reference improvements for each difficulty Program. On the basis of this research, (Pang et al., 2019) proposed a new evaluation model for the problem that it is difficult to use parallel corpus as a reference to evaluate the effect when parallel corpus is not suitable for training. Based on the commonly used style transfer accuracy measurement (Accuracy), it adds semantic retention and fluency measures. Then, for the proposed evaluation model, (Shen et al., 2017) The text style transfer model proposes a variety of improved Loss calculation methods. The results show that its Cyclic Consistency Loss has a significant effect on improving the semantic retention index. At the same time, Language Loss has a certain effect on improving the fluency index, but sometimes it sacrifices the semantic retention score.

7.2 Research status of text style transfer assessment

The evaluation methods of text style transfer can be classified from two aspects. First, according to whether there is a standard text rewritten manually by the manual, it is divided into a reference evaluation and a non-reference evaluation. Since most studies in the field of text generation do not have parallel texts, evaluation without reference is the main type of evaluation. Second, according to the evaluation index score calculation method, it is divided into two methods: manual scoring and

automatic machine calculation. The former calculates the index score automatically by writing a corresponding algorithm program, and the latter generally recruits evaluation volunteers or online crowdsourcing. (For example, recruit using the Amazon Turk platform), manually score the model's style transfer effect. In actual research, because there is no unified indicator to measure the quality of conversion results, the evaluation system in the field of text generation generally includes two parts: automatic calculation of scores and manual scores.

For the style transfer model based on parallel corpus training, researchers can compare the text after style transfer using the model with the reference text, and use the similarity (Similarity) of the two as the evaluation index of the transfer effect.

For evaluations without parallel corpora, the most widely evaluated indicators are Style Transfer Intensity and Accuracy. The two indicators focus on measuring the degree of style difference between the original text and the post-migration text: the greater the difference, the greater the style transfer intensity; the closer to the target style, the higher the style transfer accuracy.

However, if only the style change is considered as the evaluation index of the model, the migrated text will become more and more out of the original text's representation, resulting in the risk of losing the semantic information of the original text. If the text loses the meaning of the original text after migration, the application value of text style transfer is reduced. Therefore, some researchers add Content Preservation or Semantic Preservation to the evaluation index of the model. The indicator is used to measure the semantic similarity between the original text and the migrated text. The closer the two are, the more fully the meaning of the original text is expressed after the migration.

Naturalness (Naturalness) or Fluency (Fluency) is used as an evaluation indicator to measure the humanity of text in natural language generation. It is also used more and more in the evaluation of style transfer models. The closer the text is to the expression of human language, the better the text generation effect.

7.3 Calculation of text style transfer index

Under the condition that there is a reference text, when the machine calculates the score, the similarity between the migrated text and the reference text can be measured by calculating the BLEU value to reflect the migration effect (Li et al., 2018). For manual scoring, the 5-point or 10-point scale can be used to allow the evaluator to rate the similarity between the migrated text and the reference text. The three indicators can also be used to allow the evaluator to select two texts based on the style expression, content retention and fluency. The better party measures the effect of migration.

In the training of style transfer models of non-parallel corpora, for the accuracy of style transfer, machine evaluation usually uses a style classifier based on pre-training of different style texts to classify the post-migration text generated by the model. The proportion of the samples classified as the target style in the classification result is used as a measure of the accuracy of style transfer. In terms of style transfer strength, it is generally calculated using the text distance measure. For example, literature uses EMD values to evaluate style transfer strength. The larger the value, the greater the style transfer strength. In terms of manual evaluation, it is possible for humans to score the degree of manifestation of the target style of the migrated text, or to mark whether it is converted to the target style, or to score the degree of difference in style between the two.

For content retention, some studies use BLEU score instead of semantic retention, but the word change in text style transfer is sometimes deliberate, so BLEU, as an evaluation indicator of machine translation, is not very applicable to the style transfer model. Literature [10] pointed out that by comparing the similarity measurement index based on N-Gram and the similarity measurement based on word embedding, it is pointed out that for the evaluation of style transfer model, the result of using WMD value to calculate the content retention degree is closest to the result of manual evaluation. In addition, its research indicates that the effect of METEOR is also consistent with manual evaluation, but the calculation is more complicated than the former. In terms of manual evaluation, the similarity between the original text and the migrated text is generally evaluated. The scale can also be used for scoring or comparison.

In terms of naturalness, many studies reflect the degree of naturalness by calculating the perplexity of the text after migration. Among them, the literature (Pang et al., 2019) improves its calculation method, pointing out that the abnormally small confusion often manifests as a combination of vocabulary with no semantic meaning, so penalty measures are added to the calculation. Reference

also trains a classifier that can distinguish between natural and unnatural text to measure the naturalness index. In terms of manual evaluation, the naturalness of the text is often scored on a scale, which can be divided into multiple evaluation dimensions such as fluency or grammatical correctness. In addition, it is also possible for the evaluator to choose the text that he believes to be derived from human writing among the two texts without knowing whether the text is a model migration.

7.4 Discussion of style transfer assessment

7.4.1 Evaluation index

The evaluation indicators mentioned above need to consider the following issues when calculating.

For style accuracy, the scores obtained by training the style transfer classifier cannot be used as the evaluation criterion for cross-model effects, because this indicator uses a classifier trained based on different data sets, which makes horizontal comparison of different style transfer models better. When inferior, the target style score can only be used as a reference.

For content retention, especially in the application of author anonymization or writing style transfer, the concept of content retention will be very complicated. The main points of content similarity consideration can be roughly divided into the following categories: (1) corpus-specific proper nouns: the name of the same thing may be different before and after the migration, at this time the word can be marked with the NER tagger, so that these proper nouns be consistent. (2) Other content words specific to the corpus: There will be some words during the migration. Although there are many words in the corpus, they are not style keywords. In order to reduce the changes caused by the algorithm during the style transfer process, it is necessary to use the semantic similarity index to determine whether these words have been sufficiently reserved. (3) Style words: words that can reflect the text style before and after the text style transfer.

For naturalness, the common problem of using the above-mentioned common method-confusion is that the unnatural sentences may also have a lower confusion; in addition, if the frequency of content words is very small, then the content Sentences that are the same but different in style may have a high degree of confusion. There are similar problems with accuracy.

7.4.2 Score balancing and aggregation

For the three existing data: style accuracy, similarity and fluency, how to choose is also a problem. Generally speaking, these three are directly inversely proportional to each other . Different papers usually adopt different aggregation methods. Literature adopted the method of set average to calculate the comprehensive score of three indicators, and some literatures analyzed the tradeoff of each indicator score by drawing two or two indicator pairs. Therefore, the choice of expression depends on the specific use environment.

Literature believes that if a unified, cross-corpus evaluation method is needed, some sample pairs can be randomly selected, and then experts can be evaluated, and then different loss functions can be used for training. If you do not need a unified evaluation indicator, you only need to optimize the migration sentences for each indicator from the data set of interest, and iterate and optimize the algorithm continuously. This method will be more accurate when processing specific tasks more accurate.

7.5 Summary

This section summarizes the research on evaluation in the task of text style transfer, and sorts out the research status, the selection and calculation of evaluation indicators, and the problems faced by evaluation. In summary, the current evaluation indicators of this task mainly focus on three aspects of style, content and textual nature, but there is no universal calculation method in all aspects, and there is room for improvement in the calculation of each indicator. In addition, the lack of reference texts and the selection and synthesis of various evaluation indicators are the main problems facing the evaluation of migration effects.

8 Research trends and prospects of text style transfer models based on deep learning

The text style transfer model has been developed for many years since it was proposed, but with the popularization and use of deep learning in the field of natural language processing, it is still very attractive to use deep learning ideas and methods to build more efficient and accurate text style transfer models. Research direction of force. This article first introduces the basic framework of the text style transfer model, then, we organize and summarize the current public evaluation corpus and evaluation indicators used in the deep learning-based text style transfer model, and carefully compare the corpus and evaluation used by each model index.

Through the technical review of the previous chapters, it has been found that the text style transfer model based on deep learning has been vigorously developed, and the existing models are developing in a more intelligent, accurate, and general direction, which is reflected in the construction of more parallel corpora, sequences The use of the model unifies the evaluation criteria.

In summary, the success of various style transfer models is mainly due to the significant advantages of deep learning in handling complex, non-linear tasks. In essence, the existing text style transfer model cannot be deployed in practical applications. The future text style transfer model may be further developed in the following directions:

(1) Application in open and non-standard texts. Weibo, Twitter, etc. produce a large number of colloquialized, weakly regulated, and high-noise short texts. These real-time, large-scale fee-based texts have research and application value, and are widely used in sentiment analysis, social network analysis, and event monitoring, Automatic question answering and other tasks. The current text style transfer model technology is mainly in "yelp", "IMDB" and other data sets, but directly used for opening non-standard short text will inevitably lead to low performance problems. Therefore, how to produce satisfactory results for open non-normative texts is worth further exploration.

(2) Text style conversion based on the attention mechanism model. Current attention-based language models such as Bert, GPT-2, and XLNET have made breakthroughs in natural language processing tasks such as intelligent question answering, language reasoning, and emotion classification. However, at present, the research based on the attention mechanism model for text style transfer is still very rare. Used to study text style transfer. Therefore, it should be a good job to apply the language model based on the attention mechanism to the transfer of text style.

In general, although there are still some problems to be solved in the text style transfer model based on deep learning, this does not affect its further development and application in the field of natural language processing. It will still be a research for a long time in the future hot spot. The introduction of new theories, technologies and methods, especially the new achievements of natural language processing (word vector model, text distributed representation, etc.) and new models of deep learning (such as Seq2Seq, VAE, GAN) can also make them get progress.

Reference

- Kingma, Diederik P and Welling, Max. Auto-encoding variational Bayes. arXiv preprint arXiv:1312.6114, 2013.
- Miwa M, Bansal M. End-to-end relation extraction using LSTMs on sequences and tree structures Proceedings of the Annual Meeting of the Association for Computational Linguistics. Berlin, Germany, 2016: 1105-1116.
- Lin C Y, Xue N, Zhao D, et al. A convolution BiLSTM neural network model for Chinese event extraction Proceedings of the National CCF Conference on Natural Language Processing and Chinese Computing. Kunming, China, 2016: 275-287.
- Cho K, Van Merriënboer B, Gulcehre C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. Computer Science, 2014:1724-1734.
- Zhang P, Niu J, Su Z et al. End-to-end quantum-like language models with application to question answering Proceedings of the AAAI Conference on Artificial Intelligence. New Orleans, USA, 2018: 5666-5673.
- Y. Jing, Y. Yang, Z. Feng, J. Ye, Y. Yu, M. Song, "Neural style transfer: A review", IEEE Trans. Vis. Comput. Graphics.
- Reina Akama, Kazuaki Inada, Naoya Inoue, Sosuke Kobayashi, and Kentaro Inui. 2017. Generating stylistically consistent dialog responses with transfer learning. In IJCNLP.

- Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2014. Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.
- Sumit Chopra, Michael Auli, and Alexander M Rush. 2016. Abstractive sentence summarization with attentive recurrent neural networks. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 93–98.
- Sutskever, Ilya, Vinyals, Oriol, and Le, Quoc V. Sequence to sequence learning with neural networks. In Advances in neural information processing systems, pp. 3104–3112, 2014.
- Jhamtani H, Gangal V, Hovy E, et al. Shakespearizing modern language using copy-enriched sequence-to-sequence models[J]. arXiv preprint arXiv:1707.01161, 2017.
- Carlson K, Riddell A, Rockmore D. Evaluating prose style transfer with the Bible[J]. Royal Society open science, 2018, 5(10): 171920.
- Harnessing Pre-Trained Neural Networks with Rules for Formality Style Transfer, EMNLP-2019.
- Harnessing Pre-Trained Neural Networks with Rules for Formality Style Transfer, EMNLP-2019
- Semi-supervised Text Style Transfer: Cross Projection in Latent Space EMNLP 2019
- Keith Carlson, Allen Riddell, Daniel Rockmore. 2018. Zero-Shot Style Transfer in Text Using Recurrent Neural Networks
- Zichao Yang, Zhiting Hu, Chris Dyer. 2018. Unsupervised Text Style Transfer using Language Models as Discriminators
- Zhu Z, Bernhard D, Gurevych I. 2010 A monolingual tree-based translation model for sentence simplification. In Proc. of the 23rd Int.Conf. on Computational Linguistics, Beijing, China, August, pp. 1353 – 1361. Association for Computational Linguistics.
- Socher, Richard, Perelygin, Alex, Wu, Jean Y, Chuang, Jason, Manning, Christopher D, Ng, Andrew Y, Potts, Christopher, et al. Recursive deep models for semantic compositionality over sentiment treebank. In Proceedings of the conference on empirical methods in natural language processing (EMNLP), volume 1631, pp. 1642. Citeseer, 2013.
- Hofmann T. Unsupervised learning by probabilistic latent semantic analysis. Machine Learning, 2001, 42(1-2):177-196.
- Wilson, Theresa, Wiebe, Janyce, and Hoffmann, Paul. Recognizing contextual polarity in phrase-level sentiment analysis. In Proceedings of the conference on human language technology and empirical methods in natural language processing, pp. 347–354. Association for Computational Linguistics, 2005.
- Carlson K, Riddell A, Rockmore D. 2018 Evaluating prose style transfer with the Bible. R. Soc. open sci. 5: 171920.
- Mingyue Shang, Piji Li, Zhenxin Fu, Lidong Bing, Dongyan Zhao, Shuming Shi, Rui Yan. 2019. Semi-supervised Text Style Transfer: Cross Projection in Latent Space. arXiv:1909.11493v1 [cs.CL] 25 Sep 2019
- Yao Fu, Hao Zhou, Jiaze Chen, Lei Li. 2019. Rethinking Text Attribute Transfer: A Lexical Analysis arXiv:1909.12335v1 [cs.CL] 26 Sep 2019
- Mingyue Shang, Piji Li, Zhenxin Fu, Lidong Bing, Dongyan Zhao, Shuming Shi, Rui Yan. 2019. Semi-supervised Text Style Transfer: Cross Projection in Latent Space. arXiv:1909.11493v1 [cs.CL] 25 Sep 2019
- Ivan Yamshchikov, Viacheslav Shibaev and Nikolay Khlebnikov and Alexey Tikhonov. 2020. Style-transfer and Paraphrase: Looking for a Sensible Semantic Similarity Metric. arXiv:2004.05001 [cs.CL]
- Remi Mir, Bjarke Felbo, Nick Obradovich, and Iyad Rahwan. 2019. Evaluating style transfer for text. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 495–504, Minneapolis, Minnesota. Association for Computational Linguistics.
- Tianxiao Shen, Tao Lei, Regina Barzilay, and Tommi Jaakkola. 2017. Style transfer from non-parallel text by cross-alignment. In Advances in Neural Information Processing Systems 30, pages 6830–6841.
- Junbo Zhao, Yoon Kim, Kelly Zhang, Alexander Rush, and Yann LeCun. 2018. Adversarially regularized autoencoders. In Proceedings of the 35th International Conference on Machine Learning, volume 80 of Proceedings of Machine Learning Research, pages 5902–5911

- Juncen Li, Robin Jia, He He, and Percy Liang. 2018. Delete, retrieve, generate: a simple approach to sentiment and style transfer. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), pages 1865–1874. Association for Computational Linguistics.
- Richard Yuanzhe Pang. 2019. The Daunting Task of Real-World Textual Style Transfer Auto-Evaluation. arXiv:1910.03747 [cs.CL]
- Richard Yuanzhe Pang, Kevin Gimpel, 2019. Unsupervised Evaluation Metrics and Learning Criteria for Non-Parallel Textual Transfer. Association for Computational Linguistics.
- Tianxiao Shen, Tao Lei, Regina Barzilay, and Tommi Jaakkola. 2017. Style transfer from non-parallel text by cross-alignment. In Advances in Neural Information Processing Systems 30, pages 6833–6844. Curran Associates, Inc.