A GAN-based Transfer Learning Approach for Sentiment Analysis

Qian Tao†
School of Software Engineering
South China University of Technology
Guangzhou, Guangdong, P.R.China
taoqian@scut.edu.cn

Yuchen Zhou
School of Software Engineering
South China University of Technology
Guangzhou, Guangdong, P.R.China
seyczhou@gmail.com

Jie Huang
School of Software Engineering
South China University of Technology
Guangzhou, Guangdong, P.R.China
is.jiehuang@gmail.com

Jiaying Li
School of Software Engineering
South China University of Technology
Guangzhou, Guangdong, P. R. China
lijiayingjy@gmail.com

Senyu Ma
School of Software Engineering
South China University of Technology
Guangzhou, Guangdong, P. R. China
1149052711@qq.com

ABSTRACT

Transfer learning is an important artificial intelligence approach which extracts knowledge from source domain to solve tasks in the target domain. As a research hot topic, Generative Adversarial Networks (GAN) provides a powerful framework in constructing unsupervised models. A GAN consists of two neural networks: a discriminator to distinguish natural and generated samples, and a generator to deceive the discriminator. Generally, sentiment analysis of text is a big challenge in the Natural Language Processing (NLP). In this paper, A GAN-based transfer learning approach for sentiment analysis of cross-domain texts is presented. Experiments are performed in cross-domain e-commerce reviews. The results compiled demonstrate the effectiveness of the proposed approach.

CCS CONCEPTS

• **Applied computing** • Human-centered computing ~ Natural language interfaces

KEYWORDS

Transfer learning, Generative Adversarial Networks, Natural Language Processing, Sentiment Analysis, e-commerce

1 Introduction

Transfer learning is a deep learning approach to train new models on target domains by using models pretrained on source domains [1] . One major assumption of many traditional machine learning and data mining algorithms is that training and testing data are drawn from the same feature space and are of the same distribution. However, this assumption may not hold in real-world situations, since a great number of data today come from different sources. Transfer learning algorithms provide ways to extract knowledge from one or more source domains and apply the knowledge to target domains, and thus can tackle the issue. The main benefit of transfer learning is that it can solve new tasks faster as it is based on knowledge learned previously.

At present, transfer learning is widely used in a variety of fields, examples include location estimation [2], speech emotion recognition [3], sentiment analysis [4], and video annotation [5].

Despite that so much progress has been made in transfer learning researches, the existing algorithms can hardly meet the demands of the big data era [6]. Therefore, many researchers today focus on designing efficient transfer learning algorithms that can comply with the tide of big data.

The method of GAN performs well on generating real-valued data by training the generator with the guidance of a discriminator [7]. These two models seek to achieve Nash equilibrium during the iterative adversarial training process. GAN has been successfully applied to computer vision tasks like generating samples of natural images [8]. However, it is not an ideal approach to generate texts, since unlike images, texts are presented by discrete sequences, and as the discriminator can only evaluate complete sequences rather than the partial ones, the discriminator will fail to update the parameters of the generator by gradient descend. Yu et al. have introduced a model called SeqGAN to mitigate this problem [9].

Bousmalis et al. [10] have proposed a GAN-based transfer learning method on images. This model takes the rendered image and the real image as two fields, and the rendered image is corrected by GAN so that it can be similar to the real image. However, GAN-based transfer learning is still at the elementary stage of exploration and practice.

AICS 2019, July 12–13, 2019, Wuhan, Hubei, China

ACM ISBN 978-1-4503-7150-6/19/07...$15.00
https://doi.org/10.1145/3349341.3349432

Sentiment analysis has always been widely concerned and a variety of sentiment analysis algorithms has been proposed. Currently, sentiment analysis solutions can be organized into three categories: rule-based text sentiment analysis methods, machine learning-based text sentiment analysis methods and text sentiment analysis methods based on deep neural networks [11].

We introduce GAN to transfer learning to accomplish the task of cross-domain sentiment analysis in our research. Firstly, we train a text generator that produces data in the target domain from the counterparts in the source domain with GAN model, after that we use the source domain data and generated data to train a classifier to evaluate the transfer effects.

In part 2, we present the GAN-based transfer learning approach for sentiment analysis in details. Then the experiment design and results are included in part 3.

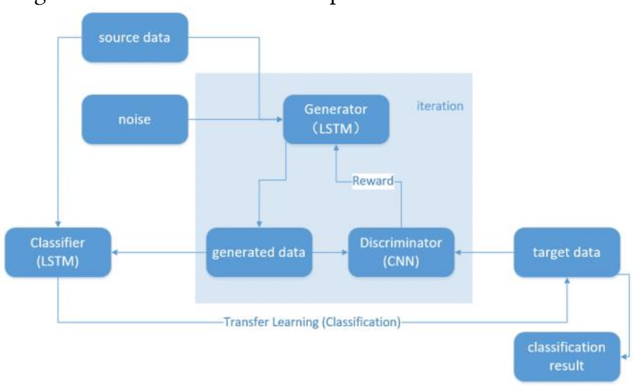


Figure 1: Flow chart of GAN-based transfer learning on texts

2 Methodology

2.1 Overall Process

Referring to previous work of transfer learning on images by Bousmalis et al. [10], we use labeled source domain data and a relatively small amount of labeled target domain data to implement the GAN-based transfer learning strategy on texts. This is inductive transfer learning in essence. Figure 1 shows the general process of the transfer learning strategy, which can be divided into two steps:

a) Transfer data from the source domain to the target domain with GAN. The labeled source domain data and the random noise are put into GAN generator, while a small number of labeled target data and the generated data are used as input of the discriminator. In the adversarial process, the expected value of rewards is calculated to update the generator so that the generated data can be more realistic. After the adversarial training, the generator can produce data which are in the same feature distribution as those in the target domain.
b) Train the classifier for specific classification tasks. The classifier is trained with source domain data as well as the generated data based on LSTM network. If we only involve the generated data in the training, multiple initializations may be needed to achieve the desired effect due to the instability of the model, whereas combining the labeled source domain data and generated text data to train the classifier is considered feasible to stabilize the model [10]. We evaluate the result of the classifier via Accuracy/Precision/Recall/F1-score so as to verify the performance of our transfer strategy.

2.2 GAN-Based Transfer

We employ a single LSTM network as the generator. It computes hidden states h_t from the embedding vectors x_t (1). Then, the hidden states are mapped to an output state distribution by a softmax output layer z as (2), where c represents the bias vector, V denotes the weight matrix. We use the convolutional neural network as the discriminator, which concatenates the embedding vectors x_t, which is T in length, to a $T \times k$ matrix $\varepsilon_{1:T}$. The matrix is convoluted with a kernel with a size of ω to obtain features (3), where ⊗ is the sum calculation of the products of elements in the direction of x_t, b is the bias term, and ρ represents a nonlinear function.

$$h_t = g(h_{t-1}, x_t) \tag{1}$$

$$p(y_t|x_1, \dots, x_t) = z(h_t) = softmax(c + Vh_t) \tag{2}$$

$$c_i = \rho(\omega \otimes \varepsilon_{1:i+l-1} + b) \tag{3}$$

The training steps are as follows:

a) Conduct model pre-training. We first pre-train the generator by minimizing the Maximum Likelihood Estimation (MLE) loss. Then we use the text data generated by the generator as the input to the discriminator and conduct pre-training on it based on the minimization of the cross entropy.
b) Start model adversarial training. In this process, the generator will perform a Monte Carlo tree search based on current tokens, namely, to produce complete sequences based on the previously generated tokens by sampling:

$$Q_{D_\phi}^{G_\theta}(s = Y_{1:t-1}, a = y) = \begin{cases} \frac{1}{N \sum_{n=1}^{N} D_\phi(Y_{1:T}^n), Y_{1:T}^n \in MC^{G\beta}(Y_{1:t};N)^{t<T}} \\ D_\phi(Y_{1:t})^{t=T} \end{cases} \tag{4}$$

The discriminator identifies the complete sequence and returns reward to the generator, enabling the parameters of the generator to be updated by Policy Gradient:

$$\Theta \leftarrow \Theta + a_h \nabla_\Theta J(\Theta) \tag{5}$$

In the meantime, the discriminator is re-trained as:

$$\min_{\Phi} -E_{Y \sim Pdata}[log D_\Phi(Y)] - E_{Y \sim G_\Theta}[\log(1 - D_\Phi(Y))] \tag{6}$$

After the adversarial process, the generator can generate data that minimize the classification correctness of the discriminator, this indicates that the generated data show the same feature distribution as source domain data. It is noteworthy that in order to map the feature distribution of the two domains into the same feature distribution and to ensure that the generated texts have a certain type of emotional polarity, the categories of the input source domain data and the target domain data need to be consistent.

2.3 Advantages of the Transfer Strategy

Compared to other traditional transfer learning methods, the transfer strategy we present has the following two advantages:

a) Uses of the generated data are flexible. Data generated by the GAN-based transfer learning can be applied to different tasks as the process of task solving is independent of data generation.

b) A large amount of target domain data can be generated for further use. By adjusting source domain data and the input noise, an unlimited number of samples can be generated for training, while this is not the case with traditional transfer learning methods.

3 Experiments

3.1 Datasets and Experimental Environment

Experiments are carried out electronic product comments as well as film and television product comments. We obtain the above datasets partly through preprocessing customer comments crawled from Amazon website and partly from public datasets of Amazon [12]. Each review in these datasets is labeled negative if it has been rated 1 or 2 stars and positive for 3 or 4 stars. We conduct the GAN-based transfer learning experiment based on the datasets and evaluate the performance with the Accuracy/Precision/Recall/F1-score indicators. The experiment runs on a Linux server with 32-core CPU and 156G memory.

3.2 Experimental Procedure

3.2.1 Fitting Experiment of the Classifier. To avoid overfitting from the overtrained classifier, fitting experiments are first performed. In the experiment, train_loss and val_loss represent the error between the predicted value and the ground truth on the training set and the validation set, and train_acc and val_acc denote the accuracy of the prediction on the training set and the validation set, respectively. The changes of the accuracy and loss values in the training process are applied to determine whether there has been an overfitting problem.

There is a total of 30000 pieces of film and television comments in the dataset, of which 24000 pieces are used as training sets, 3000 pieces as validation sets, and other 3000 pieces are for test sets. The paper [13] reported that a smaller batch_size can make the weight update more random, improve the randomness of gradient descent, and lead to better generalization. Since the size of the dataset in this experiment is about 30000 and the training time is acceptable, the batch_size is set to 64.

In this case, the correlation between the epochs and the fitting is tested and the performance is then evaluated from the following two aspects:

a) Indicators of Accuracy/Precision/Recall/F1-score under different epochs.

b) Trends in the test_loss and val_loss curves.

3.2.2 Transfer Performance Experiment. In this experiment, the classifier is trained with the source domain data and the text data generated by the GAN-based transfer. The performance of transfer is then evaluated based on the classification effects. The film and television product evaluations are used as the source domain dataset and those of electronical products as the target domain dataset. Datasets of source domain data and generated text data are divided into three groups in order to train three classifiers, numbered A, B and C shown in Table 1.

Table 1: Composition of datasets in three groups of experiment

Group	Training Set	Validation Set	Test Set
A	8000 source domain data	2000 source domain data	2000 target domain data
B	24000 source domain data and generated text data	3000 source domain data and generated text data with the ratio of 1:2	3000 target domain data
C	24000 source domain data and generated text data	6000 source domain data and generated text data with the ratio of 1:2	2000 target domain data

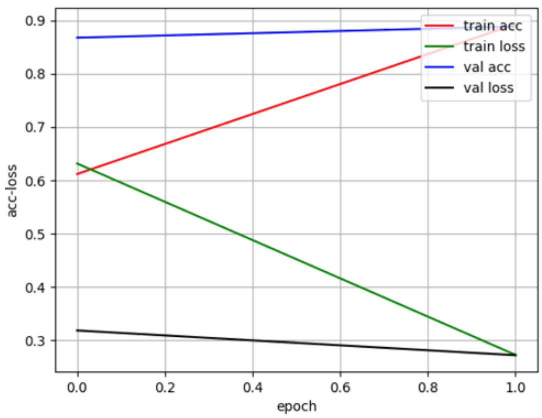
4 Results and Discussion

4.1 Fitting Experiment

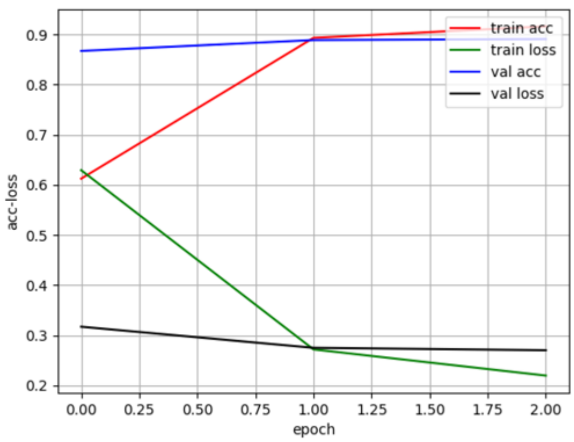
Table 2 shows the values of the four indicators under different epochs. When epoch is 3, the Accuracy, which represents the accuracy of prediction, is the highest, so are the F1-score, the average of the Precision and the Recall.

Table 2: Indicators of Accuracy/Precision/Recall/F1-score under different epochs

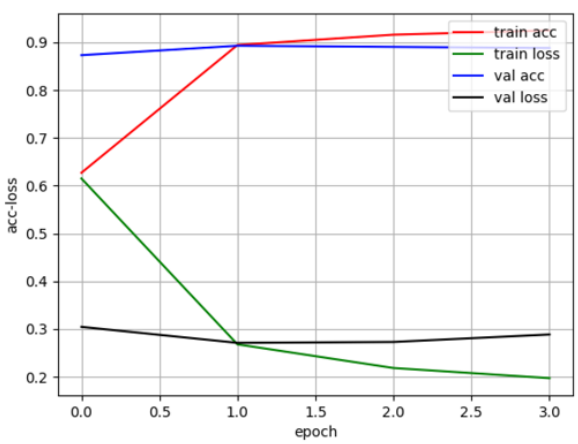
Epoch	Accuracy	Precision	Recall	F1-score
2	0.882	0.856	0.917	0.886
3	0.893	0.905	0.875	0.890
4	0.891	0.899	0.880	0.889



(a) Final epoch=2



(b) Final epoch=3



(c) Final epoch=4

Figure 2: The curves of accuracy and loss with different final epochs

Figure 2 (a - c) shows the curves of accuracy and loss with epoch when the final epoch is 2, 3, and 4, respectively.

From Figure 2 (a) we can see that, when the final epoch is 2, the curves of train_acc and val_acc have a tendency to increase, while the curves of train_loss and val_loss tend to decrease. This indicates that the model is still under-fitting and can continue to be optimized.

According to Figure 2 (b), when the final epoch is 3, the curves of val_acc and val_loss have gradually stabilized while the training accuracy still rises and the training loss still declines but at much slower rates.

Figure 2 (c) shows that while the training loss declines and training accuracy climbs even slower, the validation accuracy sees a slight drop and the validation loss increases modestly during the third training epoch. This indicates that the model is already over-fitting, as it over-learns the data in the training set, which will affect the performance of the classifier in classifications of new data.

In summary, it is found that during the process of the classifier training, the proposed model performs best when the final epoch is 3 with the batch_size of 64 and the dataset size of about 30000. In the subsequent process, the classifiers will be trained with batch_size=64 and epoch=3.

4.2 Transfer Performance Experiment

Table 3 presents five indicators of three classifiers trained with different groups of data and Figure 3 shows the comparison between them.

According to the result, all the indicators of group B have reached up to over 90%, which indicates that the classifier trained with both the generated text data and the source domain data has a good performance of classification in its own domain (which consists of the source domain data and the generated text data). Compared with group A, the indicators of group C are higher except Precision, which is slightly lower than that of group A. The comparison of indicators between group C and group A shows that the proposed method has achieved the transfer but the performance is general and needs further optimization.

Table 3: Test_loss / Accuracy / F1-score / Precision / Recall indicators of three classifiers trained with different groups of data

Group	test_loss	Accuracy	F1-score	Precison	Recall
A	0.490	0.798	0.770	0.890	0.680
B	0.183	0.934	0.935	0.919	0.952
C	0.527	0.811	0.802	0.843	0.765

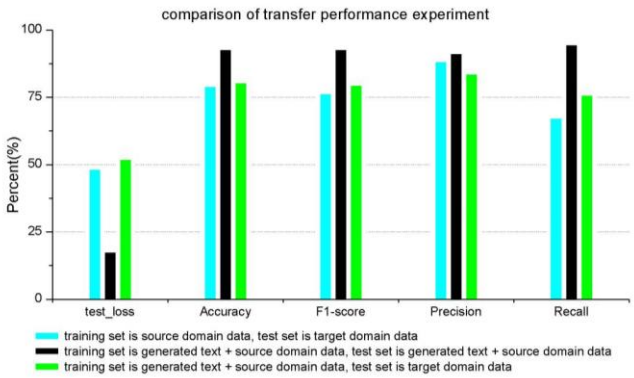


Figure 3: Comparison between five indicators of three classifiers trained with different groups of data

The following conclusions could be drawn according to the experiment: in the GAN-based transfer learning strategy, the classification performance of the training classifier on the generated text data is good, but the performance of transfer is general which needs further optimization.

5 Conclusion

The proposed method for sentiment analysis is a combination of the GAN and transfer learning. Data generated by the GAN-based transfer learning can be applied to different tasks as the processes are independent from data generation by transfer learning. The experiments were conducted on e-commerce comments have verified the validity of the proposed model. According to the experimental result, the training classifier performs well on the classification tasks of the generated text data while the performance of transfer is general and further optimization is needed for better transfer effect.

ACKNOWLEDGMENTS

This work is supported in part by the Guangdong province Natural Science Foundation under Grant No.2015A030310194, 2018A030313396, in part by the Guangzhou Major Scientific and Technological Projects under Grant No.201802010025, 201604010017, and 2019PT103, in part by the Guangdong Province Science and Technology Project under Grant No.2015A080804019, 2015B010131003.

REFERENCES

[1] L. Torrey, J. Shavlik. Transfer learning[M]//Handbook of research on machine learning applications and trends: algorithms, methods, and techniques. IGI Global, 2010: 242-264.

[2] H.S. Seok, K.B. Hwang, B.T. Zhang (2011). Feature relevance network-based transfer learning for indoor location estimation[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 41(5): 711-719.

[3] J. Deng, Z. Zhang, E. Marchi, B. Schuller (2013). Sparse autoencoder-based feature transfer learning for speech emotion recognition[C]//2013 Humaine Association Conference on Affective Computing and Intelligent Interaction. IEEE, 511-516.

[4] H.D. Zhao, B. Wu, H.Y. Wang, C. Shi (2014). Sentiment analysis based on transfer learning for Chinese ancient literature. Behavior Economic and Social Computing (BESC) International Conference on, 1-7.

[5] H. Wang, X. Wu, Y. Jia (2016). Heterogeneous domain adaptation method for video annotation[J]. IET Computer Vision, 11(2): 181-187.

[6] J. Qiu, Q. Wu, G. Ding, Y. Xu, S. Feng (2016). A survey of machine learning for big data processing[J]. EURASIP Journal on Advances in Signal Processing, (1), 67.

[7] Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio (2014). Generative adversarial nets. In NIPS.

[8] H. Huang, P.S. Yu, C. Wang (2018). An introduction to image synthesis with generative adversarial nets[J]. arXiv preprint arXiv:1803.04469.

[9] L. Yu, W. Zhang, J. Wang, Y. Yu (2017). Seqgan: Sequence generative adversarial nets with policy gradient. In AAAI.

[10] K. Bousmalis, N. Silberman, D. Dohan, D. Erhan, D. Krishnan. Unsupervised Pixel-Level Domain Adaptation with Generative Adversarial Networks. (2017). 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) - honolulu, 95-104.

[11] B. Pang, L. Lee (2008). Opinion mining and sentiment analysis[J]. Foundations and Trends® in Information Retrieval, 2(1–2), 1-135.

[12] J. McAuley, R. Pandey, J. Leskovecucehua. Inferring Networks of Substitutable and Complementary Products[J]. KDD, 2015

[13] N.S. Keskar, D. Mudigere, J. Nocedal, M. Smelyanskiy, P.T.P. Tang. On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima[J]. ICLR, 2017.