Adaptive Transfer Learning for Multi-Label Emotion Classification

Anonymous ACL submission

Abstract

In this study, we explore how data annotated with different taxonomies can be used to improve multi-label emotion classification. We propose a novel transfer learning framework to model the interaction between emotion categories, and introduce an adaptive aggregation mechanism to fuse the information from different taxonomies. The cross-taxonomy emotion interaction allows the source and target tasks to collaborate effectively, resulting in more accurate predictions. The experimental results on the SemEval-2018 dataset show that our approach can effectively boost the performance gain brought by transfer learning, and significantly outperforms existing methods.

1 Introduction

Textual emotion recognition aims to detect the emotions expressed in text. It has a wide range of applications, such as emotional chatbots (Zhou et al., 2018; Ghosal et al., 2019) and consumer analysis (Herzig et al., 2016; Alaluf and Illouz, 2019). This task is typically formalized as a multi-label emotion classification (MLEC) problem: A sentence is assigned one or more labels from a standard emotion set, such as *anger*, *disgust*, *fear*, *happiness*, *sadness*, and *surprise*.

Previous studies have focused on two approaches to improving MLEC, namely emotion association and transfer learning. Emotion association is based on the observation that emotions are interrelated (Xu et al., 2020; Alhuzali and Ananiadou, 2021). For example, *love* usually appears with *trust*, instead of *anger* or *disgust*. Thus, modeling the dependencies between emotion categories can help identify emotions more accurately. Transfer learning uses auxiliary tasks, such as sentiment classification (Liu, 2012), to facilitate the learning of MLEC (Baziotis et al., 2018; Yu et al., 2018). In fact, sentiment classification can be regarded as a special MLEC problem that contains three coarse-grained emotion categories, i.e., *positive*, *negative*, and *neutral*. In emotion analysis, researchers have proposed various taxonomies, such as the wheel of emotions created by Plutchik (Plutchik, 1980) and the six basic emotions defined by Ekman (Ekman, 1984). Datasets based on different taxonomies have also been created for different research purposes. Transfer learning makes it possible to use the data annotated with one taxonomy to improve the classification task corresponding to another taxonomy.

However, previous studies have ignored the important role of cross-taxonomy emotion interaction in transfer learning. In fact, emotions in different taxonomies are mutually indicative. For example, *anger* and *surprise* exist in both the Ekman model and the Plutchik model, and *enjoyment* in the Ekman model is closely related to *joy* and *trust* in the Plutchik model. Therefore, modeling the correspondences between emotion categories across taxonomies is expected to further enhance MLEC.

In this study, we propose an adaptive transfer learning (AdaTrans) framework for MLEC. The framework learns the correlations between emotion categories in the source and target taxonomies, and maps the probability distribution from one taxonomy to the other. Thus, the target task can utilize the output of the source task to improve its prediction, and vice versa. Moreover, we introduce an adaptive aggregation mechanism to fuse the predictions from the two taxonomies. Experimental results indicate that the cross-taxonomy emotion interaction can effectively boost the performance gain brought by transfer learning. Further analysis demonstrates the effectiveness of our proposed adaptive aggregation mechanism.

2 Related Work

For textual emotion recognition, early studies utilized emotion lexicons to discover affective words and determine their associations with emotions (Tokuhisa et al., 2008; Wen and Wan, 2014). Commonly used lexicons include WordNet-Affect

(Strapparava and Valitutti, 2004), NRC-EmoLex (Mohammad and Turney, 2013), and EmoSenticNet (Poria et al., 2014). Other studies used labeled datasets to train machine learning models for emotion classification, such as support vector machines (Liew and Turtle, 2016) and logistic regression classifiers (Park et al., 2018).

Recently, deep learning models have been applied to MLEC with promising results. Some studies have attempted to model the dependencies between emotion categories to make more accurate predictions. For example, Huang et al. (2021) used a sequential decoder to model emotion correlations implicitly. Xu et al. (2020) captured the dependencies among emotions through graph neural networks. Alhuzali and Ananiadou (2021) employed Transformers (Vaswani et al., 2017) to achieve emotion interaction.

Considering the reliance of deep learning models on large-scale labeled datasets, some studies have attempted to improve the generalization ability of neural networks through transfer learning. Baziotis et al. (2018) first pre-trained a deep learning model on a sentiment classification dataset, and then fine-tuned the model for MLEC. Yu et al. (2018) used a long short-term memory (LSTM) (Hochreiter and Schmidhuber, 1997) network to extract shared features for sentiment and emotion classification, and another LSTM network to capture emotion-specific features for MLEC. While most existing transfer learning methods focus on optimizing the feature extraction process in the encoding stage, this study is devoted to modeling the cross-taxonomy emotion interaction in the decoding stage. This allows our framework to maximize the benefits of transfer learning.

3 Approach

Suppose there are two datasets annotated with different taxonomies: $D^{\mathcal{S}} = \{\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}\}_{i=1}^{N^{\mathcal{S}}}$ for the source task and $D^{\mathcal{T}} = \{\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}\}_{i=1}^{N^{\mathcal{T}}}$ for the target task. $\boldsymbol{x}^{(i)}$ is a sentence consisting of n words, and $\boldsymbol{y}^{(i)}$ is its corresponding label set. $\boldsymbol{y}_k^{(i)} \in \{0, 1\}$ denotes whether or not $\boldsymbol{x}^{(i)}$ contains the k-th emotion in the taxonomy.

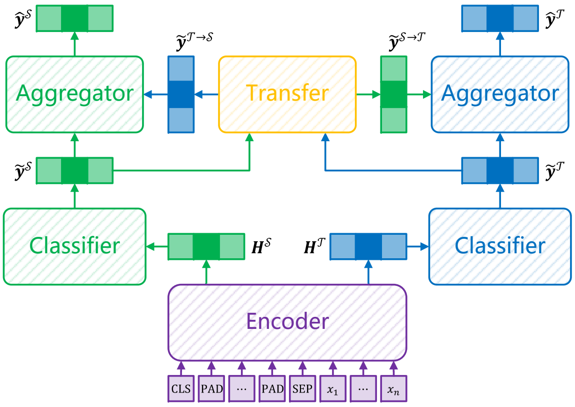

$\hat{\boldsymbol{y}}^{\mathcal{S}}$
$\tilde{\boldsymbol{y}}^{\mathcal{T}\to\mathcal{S}}$
$\tilde{\boldsymbol{y}}^{\mathcal{S}\to\mathcal{T}}$
$\hat{\boldsymbol{y}}^{\mathcal{T}}$
Aggregator
Transfer
Aggregator
$\tilde{\boldsymbol{y}}^{\mathcal{S}}$
$\tilde{\boldsymbol{y}}^{\mathcal{T}}$
Classifier
$\boldsymbol{H}^{\mathcal{S}}$
$\boldsymbol{H}^{\mathcal{T}}$
Classifier
Encoder
CLS
PAD
···
PAD
SEP
x_1
···
x_n

Figure 1: Architecture of AdaTrans.

Encoder. The overall architecture of AdaTrans is illustrated in Figure 1. Inspired by Alhuzali and Ananiadou (2021), we use BERT (Devlin et al., 2019) as an encoder, and its input is the concatenation of several placeholders and the input sentence: $[\mathtt{CLS}] + [\mathtt{PAD}] \times C^{\mathcal{S}} + [\mathtt{PAD}] \times C^{\mathcal{T}} + [\mathtt{SEP}] + \boldsymbol{x}$, where [CLS], [PAD], and [SEP] are special tokens; $C^{\mathcal{S}}$ and $C^{\mathcal{T}}$ denote the number of emotion categories in the source and target taxonomies, respectively. The hidden states $\boldsymbol{H}^{\mathcal{S}}$ and $\boldsymbol{H}^{\mathcal{T}}$ corresponding to the placeholders are used as task-specific sentence representations.[1]

Since the source and target tasks have the same decoding process, we only introduce the calculation details related to the target task below.

Classifier. The sentence representation $\boldsymbol{H}^{\mathcal{T}}$ is fed into a two-layer feed-forward network with ELU activation, followed by a sigmoid layer, to obtain the probability distribution over the emotion categories:

$$\tilde{\boldsymbol{y}}^{\mathcal{T}} = \sigma(\boldsymbol{W}_{C2}^{\mathcal{T}}\mathrm{ELU}(\boldsymbol{W}_{C1}^{\mathcal{T}}\boldsymbol{H}^{\mathcal{T}} + \boldsymbol{b}_{C1}^{\mathcal{T}}) + \boldsymbol{b}_{C2}^{\mathcal{T}}), \quad (1)$$

where $\boldsymbol{W}_{C1}^{\mathcal{T}}$, $\boldsymbol{W}_{C2}^{\mathcal{T}}$, $\boldsymbol{b}_{C1}^{\mathcal{T}}$, and $\boldsymbol{b}_{C2}^{\mathcal{T}}$ are learnable parameters.

Transfer. To learn the correlations between emotion categories in the source and target taxonomies, we use the probability distribution of the source task to predict the probability distribution of the target task:

$$\tilde{\boldsymbol{y}}^{\mathcal{S}\to\mathcal{T}} = \sigma(\boldsymbol{W}_{T}^{\mathcal{S}\to\mathcal{T}}\tilde{\boldsymbol{y}}^{\mathcal{S}} + \boldsymbol{b}_{T}^{\mathcal{S}\to\mathcal{T}}), \quad (2)$$

where $\boldsymbol{W}_{T}^{\mathcal{S}\to\mathcal{T}}$ and $\boldsymbol{b}_{T}^{\mathcal{S}\to\mathcal{T}}$ are learnable parameters.

Aggregator. In order to fuse the original prediction $\tilde{\boldsymbol{y}}^{\mathcal{T}}$ and the transferred prediction $\tilde{\boldsymbol{y}}^{\mathcal{S}\to\mathcal{T}}$, a weight vector is used to control the contribution of each part to the final probability distribution. The weight vector is determined dynamically during the inference process:

$$\boldsymbol{\alpha}^{\mathcal{S}\to\mathcal{T}} = \sigma(\boldsymbol{W}_{A}^{\mathcal{S}\to\mathcal{T}}[\tilde{\boldsymbol{y}}^{\mathcal{T}}; \tilde{\boldsymbol{y}}^{\mathcal{S}\to\mathcal{T}}] + \boldsymbol{b}_{A}^{\mathcal{S}\to\mathcal{T}}), \quad (3)$$

[1] We have omitted the calculation details here due to space limitations. Readers can refer to Alhuzali and Ananiadou (2021) and Devlin et al. (2019) for more information.

$$\hat{\boldsymbol{y}}^{\mathcal{T}} = (\mathbf{1} - \boldsymbol{\alpha}^{\mathcal{S}\to\mathcal{T}}) \odot \tilde{\boldsymbol{y}}^{\mathcal{T}} + \boldsymbol{\alpha}^{\mathcal{S}\to\mathcal{T}} \odot \tilde{\boldsymbol{y}}^{\mathcal{S}\to\mathcal{T}}, \quad (4)$$

where $\boldsymbol{W}_A^{\mathcal{S}\to\mathcal{T}}$ and $\boldsymbol{b}_A^{\mathcal{S}\to\mathcal{T}}$ are learnable parameters; $\odot$ denotes element-wise multiplication.

Training. The predicted probability distribution $\hat{\boldsymbol{y}}^{\mathcal{T}}$ is compared with the ground-truth label set $\boldsymbol{y}^{\mathcal{T}}$, to obtain the binary cross-entropy (BCE) loss:

$$\mathcal{L}_{\text{BCE}}^{\mathcal{T}} = -\frac{1}{C^{\mathcal{T}}} \sum_{k=1}^{C^{\mathcal{T}}} [\boldsymbol{y}_k^{\mathcal{T}} \log(\hat{\boldsymbol{y}}_k^{\mathcal{T}}) + (1 - \boldsymbol{y}_k^{\mathcal{T}}) \log(1 - \hat{\boldsymbol{y}}_k^{\mathcal{T}})]. \quad (5)$$

Following Alhuzali and Ananiadou (2021), we also employ the label-correlation aware (LCA) loss to maximize the distance between positive and negative labels:

$$\mathcal{L}_{\text{LCA}}^{\mathcal{T}} = \frac{1}{|\boldsymbol{y}^1||\boldsymbol{y}^0|} \sum_{(p,q)\in\boldsymbol{y}^1\times\boldsymbol{y}^0} \exp(\hat{\boldsymbol{y}}_q^{\mathcal{T}} - \hat{\boldsymbol{y}}_p^{\mathcal{T}}), \quad (6)$$

where $\boldsymbol{y}^1$ and $\boldsymbol{y}^0$ denote the set of positive and negative labels, respectively. The overall loss function is defined as follows:

$$\mathcal{L}^{\mathcal{T}} = (1 - \lambda)\mathcal{L}_{\text{BCE}}^{\mathcal{T}} + \lambda\mathcal{L}_{\text{LCA}}^{\mathcal{T}} + \mu||\Theta||^2, \quad (7)$$

where λ is a hyperparameter used to control the effect of the BCE loss and the LCA loss; μ denotes the coefficient of the L_2 regularization term $||\Theta||^2$.

4 Experiments

4.1 Experimental Settings

Datasets. SemEval-2018 (Mohammad et al., 2018) was used as the target dataset to evaluate our approach. It contains English tweets with 11 emotion categories (SemEval taxonomy). GoEmotions (Demszky et al., 2020) was used as the source dataset. It contains English Reddit comments annotated with three different taxonomies: Ekman (6 emotion categories), GoEmotions (27 emotion categories), and Sentiment (3 emotion categories). The statistics of the datasets are shown in Appendix A.

Metrics. Following Mohammad et al. (2018), we used Jaccard index, micro-averaged F1-score, and macro-averaged F1-score as the evaluation metrics. We repeated each experiment 10 times, and reported the average results.

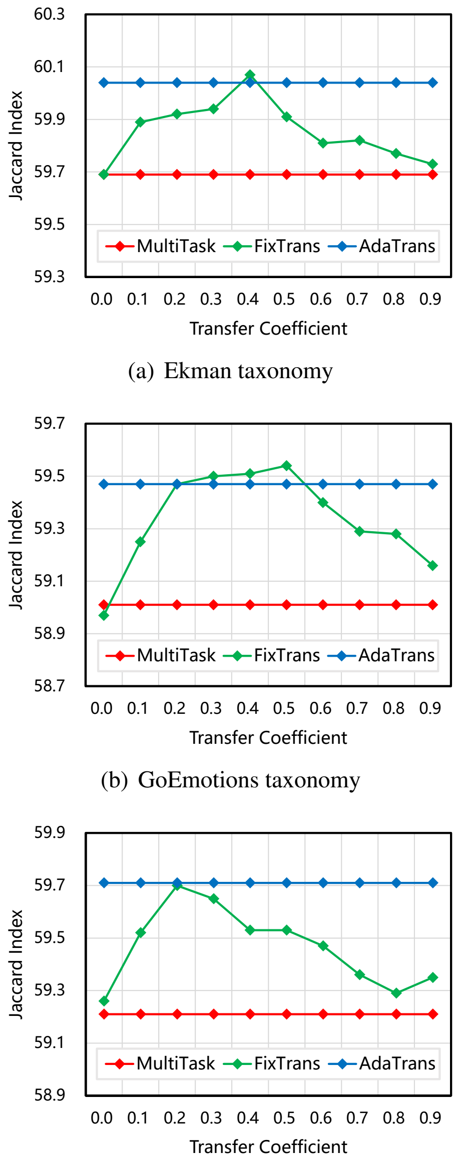
Compared Methods. PlusEmo2Vec (Park et al., 2018), TCS-Research (Meisheri and Dey, 2018), and NTUA-SLP (Baziotis et al., 2018) are the top-3 systems in the SemEval-2018 competition. Seq2Emo (Huang et al., 2021), LEM (Fei et al., 2020), BERT-GAT (Xu et al., 2020), BERT-GCN (Xu et al., 2020), and SpanEmo (Alhuzali and Ananiadou, 2021) are deep learning methods that model emotion correlations through sequential decoders, variational autoencoders, graph attention networks, graph convolutional networks, and Transformers, respectively. NTUA-SLP (Baziotis et al., 2018) and DATN (Yu et al., 2018) are transfer learning methods, based on model pre-training and attention networks, respectively. MultiTask is a variant of AdaTrans that removes the transfer and aggregation modules.

Methods	Jaccard	Micro-F	Macro-F
PlusEmo2Vec$^\natural$	57.60	69.20	49.70
TCS-Research$^\natural$	58.20	69.30	53.00
NTUA-SLP$^\natural$	58.80	70.10	52.80
Seq2Emo$^\natural$	58.67	70.02	51.92
LEM$^\natural$	-	67.50	56.70
DATN$^\natural$	58.30	-	54.40
BERT-GAT$^\natural$	58.30	69.90	56.90
BERT-GCN$^\natural$	58.90	70.70	56.30
SpanEmo†	58.60	70.71	55.58
MultiTask-Ekman†	59.69	71.18	56.73
MultiTask-GoEmotions†	59.01	70.85	55.64
MultiTask-Sentiment†	59.21	70.88	56.27
AdaTrans-Ekman†	**60.04**	**71.62**	**57.14**
AdaTrans-GoEmotions†	59.47	71.14	56.64
AdaTrans-Sentiment†	59.71	71.19	56.46

Table 1: Performance comparison of different methods. $\natural$ denotes the results retrieved from the original papers. $\dagger$ denotes the results obtained by our implementations.

4.2 Experimental Results

Table 1 shows the experimental results of different methods. SpanEmo and MultiTask have the same structure, but the former is trained only on the target dataset, while the latter also learns from the source dataset. Compared with SpanEmo, MultiTask has an improvement of 0.41% to 1.09% in terms of Jaccard index. This suggests that, although the source and target tasks possess different taxonomies, the knowledge learned from the source task can still improve the performance of the target task. Compared with MultiTask, AdaTrans achieves a Jaccard index improvement of 0.35% to 0.50% with the same training datasets. This indicates that the cross-taxonomy emotion interaction can effectively boost the performance gain brought by transfer learning. In addition, we observed that the source dataset annotated with different taxonomies contributed differently to the target task. AdaTrans-Ekman achieves the best results, and outperforms AdaTrans-GoEmotions and AdaTrans-Sentiment



(a) Ekman taxonomy

(b) GoEmotions taxonomy

(c) Sentiment taxonomy

Figure 2: Performance comparison of model variants.

by 0.57% and 0.33% respectively in Jaccard index. We believe this is because the Ekman taxonomy is more similar to the SemEval taxonomy, and therefore it is easier to learn their associations.

4.3 Analysis and Discussion

To verify the effectiveness of the adaptive aggregation mechanism in AdaTrans, we compared it with MultiTask and another variant, namely FixTrans. FixTrans uses a fixed weight to fuse the original and transferred predictions. That is, the weight vector in Equation 4 becomes a pre-defined hyperparameter (transfer coefficient). Figure 2 shows the Jaccard index of FixTrans with different transfer coefficients. We found that FixTrans performs well with a suitable transfer coefficient. However, when the coefficient is too small or too large, its performance decreases significantly. Moreover, the optimal coefficients are different for datasets annotated with different taxonomies, which makes it more difficult to determine. In contrast, AdaTrans achieves competitive results without the need to set this parameter in advance. This advantage allows AdaTrans to be used flexibly with various datasets.

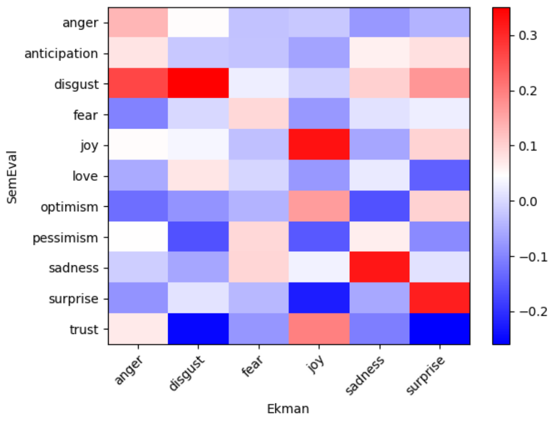


Figure 3: Emotion correlations (SemEval–Ekman).

In AdaTrans, the transfer module acts as a bridge between the source taxonomy and the target taxonomy. The association of the two taxonomies can be reflected by the mapping matrix in Equation 2. Figure 3 shows the correlations between the emotion categories in the SemEval and Ekman taxonomies. We observed high correlations between the emotions shared by the two taxonomies, such as *disgust*, *joy*, *sadness*, and *surprise*. For some unique but highly correlated emotions, such as *optimism* in SemEval and *joy* in Ekman, AdaTrans can also find their associations. Thus, our framework can not only be used for MLEC, but also provides an empirical method to reveal the intrinsic connections between different emotion taxonomies.

5 Conclusion

In this study, we propose an adaptive transfer learning framework that uses data annotated with different taxonomies to improve MLEC. The framework learns the correlations between emotion categories across taxonomies, and fuses the predictions from different taxonomies through an adaptive aggregation mechanism. The experimental results show that our method achieves state-of-the-art results on the SemEval-2018 dataset. Further analysis demonstrates the effectiveness of our approach.

References

Yaara Benger Alaluf and Eva Illouz. 2019. Emotions in consumer studies. *The Oxford Handbook of Consumption*, pages 239–240.

Hassan Alhuzali and Sophia Ananiadou. 2021. Spanemo: Casting multi-label emotion classification as span-prediction. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics*, pages 1573–1584.

Christos Baziotis, Athanasiou Nikolaos, Alexandra Chronopoulou, Athanasia Kolovou, Georgios Paraskevopoulos, Nikolaos Ellinas, Shrikanth Narayanan, and Alexandros Potamianos. 2018. Ntua-slp at semeval-2018 task 1: Predicting affective content in tweets with deep attentive rnns and transfer learning. In *Proceedings of The 12th International Workshop on Semantic Evaluation*, pages 245–255.

Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade, and Sujith Ravi. 2020. Goemotions: A dataset of fine-grained emotions. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4040–4054.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4171–4186.

Paul Ekman. 1984. Expression and the nature of emotion. In *Approaches to Emotion*, pages 319–344. Psychology Press.

Hao Fei, Yue Zhang, Yafeng Ren, and Donghong Ji. 2020. Latent emotion memory for multi-label emotion classification. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence*, pages 7692–7699.

Deepanway Ghosal, Navonil Majumder, Soujanya Poria, Niyati Chhaya, and Alexander Gelbukh. 2019. Dialoguegcn: A graph convolutional neural network for emotion recognition in conversation. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, pages 154–164.

Jonathan Herzig, Guy Feigenblat, Michal Shmueli-Scheuer, David Konopnicki, and Anat Rafaeli. 2016. Predicting customer satisfaction in customer support conversations in social media using affective features. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*, pages 115–119.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation*, pages 1735–1780.

Chenyang Huang, Amine Trabelsi, Xuebin Qin, Nawshad Farruque, Lili Mou, and Osmar R Zaiane. 2021. Seq2emo: A sequence to multi-label emotion classification model. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4717–4724.

Jasy Suet Yan Liew and Howard R Turtle. 2016. Exploring fine-grained emotion detection in tweets. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 73–80.

Bing Liu. 2012. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, pages 1–167.

Hardik Meisheri and Lipika Dey. 2018. Tcs research at semeval-2018 task 1: Learning robust representations using multi-attention architecture. In *Proceedings of the 12th International Workshop on Semantic Evaluation*, pages 291–299.

Saif Mohammad, Felipe Bravo-Marquez, Mohammad Salameh, and Svetlana Kiritchenko. 2018. Semeval-2018 task 1: Affect in tweets. In *Proceedings of the 12th International Workshop on Semantic Evaluation*, pages 1–17.

Saif M Mohammad and Peter D Turney. 2013. Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, pages 436–465.

Ji Ho Park, Peng Xu, and Pascale Fung. 2018. Plusemo2vec at semeval-2018 task 1: Exploiting emotion knowledge from emoji and # hashtags. In *Proceedings of the 12th International Workshop on Semantic Evaluation*, pages 264–272.

Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. In *Theories of Emotion*, pages 3–33. Academic Press.

Soujanya Poria, Alexander Gelbukh, Erik Cambria, Amir Hussain, and Guang-Bin Huang. 2014. Emosenticspace: A novel framework for affective common-sense reasoning. *Knowledge-Based Systems*, pages 108–123.

Carlo Strapparava and Alessandro Valitutti. 2004. Wordnet-affect: An affective extension of wordnet. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation*, pages 1083–1086.

Ryoko Tokuhisa, Kentaro Inui, and Yuji Matsumoto. 2008. Emotion classification using massive examples extracted from the web. In *Proceedings of the 22nd International Conference on Computational Linguistics*, pages 881–888.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*, pages 5998–6008.

Shiyang Wen and Xiaojun Wan. 2014. Emotion classification in microblog texts using class sequential rules. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, pages 187–193.

Peng Xu, Zihan Liu, Genta Indra Winata, Zhaojiang Lin, and Pascale Fung. 2020. Emograph: Capturing emotion correlations using graph networks. *arXiv preprint arXiv:2008.09378*.

Jianfei Yu, Luis Marujo, Jing Jiang, Pradeep Karuturi, and William Brendel. 2018. Improving multi-label emotion classification via sentiment classification with dual attention transfer network. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1097–1102.

Hao Zhou, Minlie Huang, Tianyang Zhang, Xiaoyan Zhu, and Bing Liu. 2018. Emotional chatting machine: Emotional conversation generation with internal and external memory. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, pages 730–739.

A Dataset Statistics

Table 2 and Table 3 present the statistics of the SemEval-2018 and GoEmotions datasets, respectively.

B Implementation Details

We utilized Ekphrasis[2] for data pre-processing. It is a text processing tool geared towards text from

[2] https://pypi.org/project/ekphrasis/

Training (#)	6,838
Validation (#)	886
Test (#)	3,259
Total (#)	10,983
Categories (#)	11
- Anger (%)	36.06
- Anticipation (%)	13.90
- Disgust (%)	36.60
- Fear (%)	16.83
- Joy (%)	39.32
- Love (%)	12.27
- Optimism (%)	31.27
- Pessimism (%)	11.56
- Sadness (%)	29.44
- Surprise (%)	05.15
- Trust (%)	05.04

Table 2: Statistics of the SemEval-2018 dataset.

Total (#)	38,242
Taxonomy	Ekman
Categories (#)	6
- Anger (%)	18.36
- Disgust (%)	02.65
- Fear (%)	02.43
- Joy (%)	56.83
- Sadness (%)	10.54
- Surprise (%)	17.44
Taxonomy	GoEmotions
Categories (#)	27
- Admiration (%)	13.39
- Amusement (%)	07.57
- Anger (%)	05.13
- Annoyance (%)	08.09
- Approval (%)	09.64
- Caring (%)	03.60
- Confusion (%)	04.37
- Curiosity (%)	07.12
- Desire (%)	02.09
- Disappointment (%)	04.14
- Disapproval (%)	06.75
- Disgust (%)	02.65
- Embarrassment (%)	00.98
- Excitement (%)	02.75
- Fear (%)	02.00
- Gratitude (%)	08.82
- Grief (%)	00.25
- Joy (%)	04.67
- Love (%)	06.74
- Nervousness (%)	00.54
- Optimism (%)	05.17
- Pride (%)	00.37
- Realization (%)	03.61
- Relief (%)	00.48
- Remorse (%)	01.75
- Sadness (%)	04.25
- Surprise (%)	03.48
Taxonomy	Sentiment
Categories (#)	3
- Ambiguous (%)	17.44
- Negative (%)	32.27
- Positive (%)	56.83

Table 3: Statistics of the GoEmotions dataset.

Methods	Jaccard	Micro-F	Macro-F
MultiTask-Ekman	60.37	71.77	58.36
MultiTask-GoEmotions	60.13	71.67	57.74
MultiTask-Sentiment	60.23	71.78	58.48
AdaTrans-Ekman	60.90	72.19	58.60
AdaTrans-GoEmotions	60.25	71.83	58.65
AdaTrans-Sentiment	60.16	71.62	58.49

Table 4: Results of MultiTask and AdaTrans on the validation datasets.

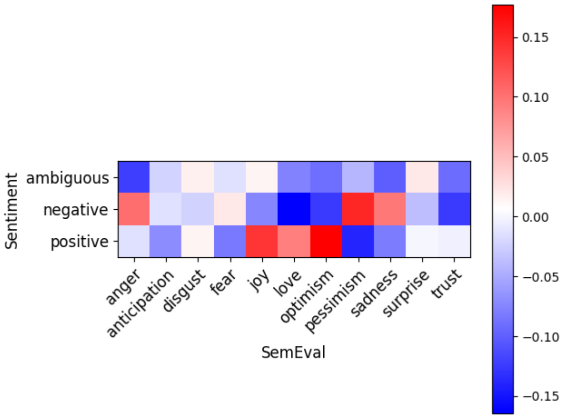


Figure 4: Emotion correlations (SemEval–Sentiment).

social networks. We used the tool for tokenization, spell correction, and word normalization.

Our framework was implemented in PyTorch[3], and trained on NVIDIA GeForce RTX 2080 Ti GPUs. We used the uncased version of BERT$_{\text{base}}$ model[4] as the encoder of AdaTrans. The dimension of hidden states was 768. The maximum input sequence length was limited to 100. The hidden size of the classifiers was set to 768. The hyperparameters λ and μ in the loss function were set to 0.2 and 1e-5, respectively.

For model training, we sampled the mini-batch alternately from $D^{\mathcal{S}}$ and $D^{\mathcal{T}}$. The batch size was set to 32. We used the BERTAdam optimizer to update the model parameters. The initial learning rate was set to 2e-5 and 1e-3 for fine-tuning BERT and optimizing other modules, respectively. We trained the model for 20 epochs, and adopted a linear learning rate decay schedule. The best model was selected based on the Jaccard index on the validation set. To avoid overfitting, we performed early stopping with a patience of 5.

C Results on Validation Datasets

The experimental results of MultiTask and AdaTrans on the validation datasets are shown in Table 4.

D More Visualizations

Figure 4 shows the correlations between the emotion categories in the SemEval and Sentiment taxonomies. Notably, *anger*, *pessimism*, and *sadness* in SemEval are closely related to *negative* in Sentiment. Meanwhile, *joy*, *love*, and *optimism* in SemEval are more related to *positive* in Sentiment.

Figure 5 shows the correlations between the emotion categories in the SemEval and GoEmotions taxonomies. It can be found that *joy* in SemEval is closely related to *admiration* and *amusement* in GoEmotions. In addition, *optimism* in SemEval is highly correlated with *approval* and *caring* in GoEmotions.

[3] https://pytorch.org

[4] https://huggingface.co/bert-base-uncased

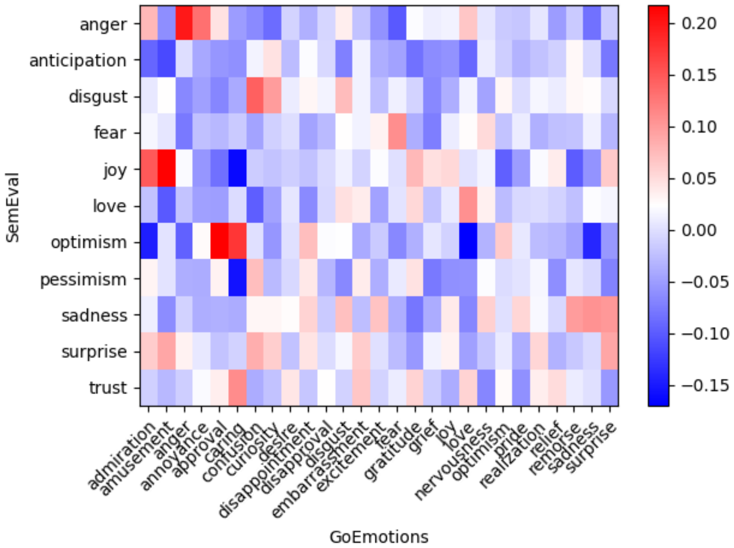


Figure 5: Emotion correlations (SemEval–GoEmotions).