Cross-Domain Sentiment Classification Using Fine-Tuned Transformer Models

Badabagni Rohini
rohini259@gmail.com

***Abstract*—Cross-domain sentiment classification builds a sentiment model that generalizes across different domains despite shifts in vocabulary and context. Traditional domain adaptation methods depend heavily on manual features and struggle to scale across diverse domains. In this work, we study the effectiveness of pretrained transformer-based models for cross-domain sentiment classification. Specifically, we fine-tune BERT and XLNet on a labeled source domain and evaluate their performance on unseen target domains using the Amazon multi-domain sentiment dataset. Experimental results demonstrate that transformer-based models achieve strong cross-domain generalization, with XLNet consistently outperforming BERT.**

***Index Terms*—Sentiment Analysis, Cross-Domain Classification, Transfer Learning, BERT, XLNet**

I. Introduction

Sentiment analysis is a fundamental task in natural language processing (NLP). It focuses on identifying opinions and emotions from textual data. It plays a critical role in applications such as customer feedback analysis, opinion mining, and social media monitoring. Most traditional sentiment classification systems assume that the training and testing data come from the same domain. However, in real-world scenarios, labeled data is often available only for specific domains. The models are expected to perform well on new and unseen domains. This mismatch leads to what is known as domain shift, where performance drops sharply.

Models like BERT and XLNet, trained on massive text can be fine-tuned for specific tasks. These models enable transfer learning through fine-tuning. Recent work by Myagmar *et al.* [1] demonstrated that fine-tuned BERT and XLNet significantly perform well than traditional methods for cross-domain sentiment classification.

A. Objectives

The objective of this project is to evaluate pretrained transformer models for cross-domain sentiment classification. The objectives are:

- To implement a cross-domain sentiment classification framework using a dataset [2].
- To apply transfer learning via fine-tuning using BERT and XLNet.
- To evaluate performance across multiple source–target domain pairs.
- To compare BERT and XLNet in cross-domain settings.

II. Literature Review

Myagmar *et al.* [1] were the first to systematically investigate bidirectional contextualized transformer language models for cross-domain sentiment classification. Their work showed that fine-tuned BERT and XLNet models achieve state-of-the-art performance on the Amazon multi-domain sentiment dataset while using substantially fewer labeled samples than traditional pivot-based and adversarial domain adaptation approaches. Cross-domain sentiment classification has been extensively studied in NLP. Blitzer et al. [2] introduced Structural Correspondence Learning (SCL), which aligns domain-specific features using pivot words and established the Amazon multi-domain sentiment dataset as a benchmark. This dataset remains widely used in domain adaptation research.

The transformer architectures in NLP, BERT [3] introduced bidirectional self-attention and showed that pretraining on large text, fine-tuning. XLNet [4] built on this idea with a permutation-based autoregressive objective enabling it to capture bidirectional context without relying on masking.

III. Methodology

A. System Design

The proposed system follows a cross-domain training and evaluation framework. A pretrained transformer model is fine-tuned using labeled data from a source domain and evaluated on a different target domain without access to target-domain labels during training. The experimental protocol and fine-tuning strategy adopted in this study are inspired by the framework proposed in [1].

B. Tools and Models

Our implementation was in Python using PyTorch, HuggingFace Transformers library. We focused on two pretrained models: BERT-base (cased) and XLNet-base (cased) with their pretrained weights initialized and then fine-tuned on the source domain data. Transfer learning allows the models to adapt their general language understanding to the specific task of sentiment classification across domains.

C. Evaluation

Accuracy was measured on unseen target domains during training. Experiments were conducted across multiple source–target domain pairs and the results were summarized in tabular form to provide a clear comparison between BERT and XLNet across different domain shifts.

IV. Implementation

The code was written in Colab Notebook with T4-GPU. The dataset was uploaded to and accessed from Google Drive. Transformer tokenizers handle raw text effectively so minimal preprocessing was done. Both BERT and XLNet models were fine-tuned. Training was carried out for three epochs, which provided a good balance preventing overfitting.

V. Results

Experiments confirm that both BERT and XLNet perform well in cross-domain sentiment classification. The observed superiority of XLNet over BERT is consistent with the findings reported in [1].

VI. Conclusion

In summary, our study shows that pretrained transformers are effective for cross-domain sentiment classification. Fine-tuning allows them to generalize well. Among the models tested, XLNet consistently outperformed BERT, highlighting its strength in handling domain shifts.

A. Limitations

- The study was limited to a relatively small dataset.
- No hyperparameter tuning techniques were explored.

B. Future Work

Future research could explore larger transformer variants, multi-source domain training, domain-adversarial learning, and hyperparameter optimization to further improve performance.

References

[1] B. Myagmar, J. Li, and S. Kimura, "Cross-domain sentiment classification with bidirectional contextualized transformer language models," *IEEE Access*, vol. 8, pp. 163995–164011, 2020, doi: 10.1109/ACCESS.2020.3021571.

[2] J. Blitzer, M. Dredze, and F. Pereira, "Domain adaptation for sentiment classification," *Proc. ACL*, pp. 440–447, 2007.

[3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *Proc. NAACL-HLT*, 2019.

[4] Z. Yang *et al.*, "XLNet: Generalized autoregressive pretraining for language understanding," *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.