AI6122- Directed Reading

Maheswaran RohinKumar - G2303513K
rohinkum001@e.ntu.edu.sg

ABSTRACT

This directed reading report examines three papers from recent SIGIR conferences to delve into the field of sentiment analysis and sentiment classification. The chosen papers,"Tag-assisted Multimodal Sentiment Analysis under Uncertain Missing Modalities,[1]", "Aspect Feature Distillation and Enhancement Network for Aspect-based Sentiment Analysis,[2]" and "Iterative Network Pruning with Uncertainty Regularization for Lifelong Sentiment Classification,[3]". The report provides a thorough overview of the developments, approaches, and difficulties in Sentiment Analysis and Classification by synthesizing the most important findings from these papers.

ACM Reference Format:
Maheswaran RohinKumar - G2303513K. 2018. AI6122- Directed Reading. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX).* ACM, New York, NY, USA, 3 pages. https://doi.org/XXXXXXX.XXXXXXX

1 PAPER TITLE : TAG-ASSISTED MULTIMODAL SENTIMENT ANALYSIS UNDER UNCERTAIN MISSING MODALITIES

The term "missing modalities" in multimodal sentiment analysis describes situations in which particular modalities like audio, visual, or textual data—are restricted or unavailable, producing insufficient input data for sentiment analysis.The Tag-Assisted Transformer Encoder (TATE) network is introduced in this paper as a potential remedy for multimodal sentiment analysis's issue of missing ambiguous modalities.The proposed model addresses the challenge of missing modalities by incorporating a tag encoding module and a common space projection pattern. It utilizes a Transformer encoder-decoder network for learning missing modality features and applies the learned representations for sentiment classification.

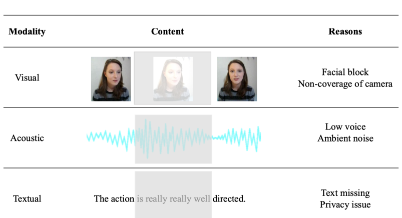

Figure 1: Examples of missing modalities.[1]

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

The Dataset that has been used in this paper are CMU-MOSI and IEMO-CAP datasets.The CMU-MOSI dataset is a multimodal collection of YouTube video clips with sentiment labels annotated on each clip. This dataset includes textual (transcripts), acoustic (speech), and visual (facial expressions) modalities. Another multimodal dataset with audiovisual recordings of actors portraying a range of emotions is called IEMO-CAP. It covers emotions like happiness, sadness, anger, frustration, excitement, and surprise and incorporates modalities like text, audio, and visual information. The TATE network consists of several key components and steps, which are crucial for handling missing modalities and improving sentiment analysis :

Tag Encoding Module: The Tag Encoding Module in the TATE network is vital for handling missing modalities. It marks and identifies gaps, directing attention to specific missing information. By using encoded tags, the TATE network adapts its processing, ensuring robust performance in the absence of certain modalities. It not only flags missing modalities but actively contributes to accurate predictions based on available information.

Common Space Projection Module: The Common Space Projection Module in TATE enhances joint representations by mapping intra-modal features into a shared space. Using a two-by-two projection pattern and linear transformations, it efficiently integrates data from different modalities, creating self-related common spaces. This ensures robust sentiment analysis, even with missing modalities.

Transformer Encoder: In TATE, the Transformer encoder is crucial for multimodal sentiment analysis. It captures intra-modal features, enabling long-term dependency modeling. This essential component processes input data, creates detailed representations, and guides feature reconstruction for missing modalities. The result is improved sentiment classification accuracy and enhanced capability to handle uncertainty.

Comparison with Baselines: The paper presents a thorough comparison of the TATE network with various baselines, including state-of-the-art models, to evaluate its performance in handling missing modalities and enhancing sentiment analysis accuracy.The results of all baselines for missing a single modality are provided, showcasing the performance of different models across different levels of missing modalities (0, 0.1, 0.2, 0.3, 0.4, and 0.5). Higher M-F1 and ACC scores across the CMU-MOSI and IEMOCAP datasets consistently show that the TATE network performs better than the baselines. Table 3 in the paper reveals TATE's superiority in handling missing multiple modalities, achieving higher M-F1 and ACC scores across various scenarios on both CMU-MOSI and IEMOCAP datasets. The paper emphasizes TATE's robustness, showcasing improved scores even when multiple modalities are randomly discarded. Overall, TATE consistently outperforms baselines, demonstrating its effectiveness and resilience in addressing the challenge of missing modalities in multimodal sentiment analysis.

Experimental Results: The TATE network consistently outperforms baselines in addressing missing modalities in multimodal sentiment analysis. It shows superior M-F1 and ACC scores on CMU-MOSI and IEMO-CAP datasets, even in cases of single or multiple missing modalities. Experimental results highlight TATE's effectiveness across various settings, solidifying its performance advantage over baselines.

2 PAPER TITLE : ASPECT FEATURE DISTILLATION AND ENHANCEMENT NETWORK FOR ASPECT-BASED SENTIMENT ANALYSIS

The paper proposes an Aspect Feature Distillation and Enhancement Network (AFDEN) for aspect-based sentiment analysis. The network aims to extract aspect-related and aspect-unrelated features, eliminate interference from aspect-unrelated features, and enhance aspect-related features using supervised contrastive learning. The proposed model achieves state-of-the-art performance on benchmark datasets and demonstrates effectiveness and robustness.

The datasets used in this paper are the Restaurant and Laptop datasets from SemEval2014 task 4 , the Twitter dataset originally built by , and the Multi-Aspect Multi-Sentiment (MAMS) dataset . Each dataset contains reviews with sentiment polarities annotated for aspects within the sentences. The statistics for these datasets are provided in Table 1 of the paper

The overall steps carried out in the proposed Aspect Feature Distillation and Enhancement Network (AFDEN) for aspect-based sentiment analysis are as follows:

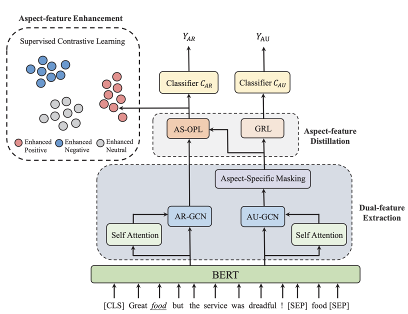

Figure 2: The overall architecture of AFDEN.[2]

Dual-Feature Extraction: To extract aspect-related and aspect-unrelated features from the input sentence-aspect pairs, the model makes use of a dual-feature extraction module.
Aspect-Feature Distillation: : A novel aspect-feature distillation module is employed, which includes a gradient reverse layer (GRL) and an aspect-specific orthogonal projection layer (AS-OPL). The GRL learns aspect-unrelated contextual features through adversarial training, while the AS-OPL further projects aspect-related features into the orthogonal space of aspect-unrelated features
Aspect-Feature Enhancement:A module for improving aspect-features is created, utilizing supervised contrastive learning to identify implicit information both within and between sentiment labels. The sentiment discriminability of aspect features is improved by this module.

Current state-of-the-art (SOTA) models in the field of aspect-based sentiment analysis have been compared with the proposed Aspect Feature Distillation and Enhancement Network (AFDEN). The study shows that AFDEN outperforms multiple baseline models and reaches state-of-the-art performance on benchmark datasets. To predict sentence-aspect pairs, the AFDEN model employs a dual-feature extraction module, aspect-feature distillation, and aspect-feature enhancement. To extract aspect-related and aspect-unrelated features, it combines self-attention mechanisms, graph convolutional networks, and aspect-specific masking. The model's effectiveness and robustness are verified through ablation studies and case studies, further supporting its superiority over existing approaches.
Experimental Results: The results of the experiment show that the suggested Aspect Feature Distillation and Enhancement Network (AFDEN) is a useful tool for aspect-based sentiment analysis.Across several datasets, ablation studies demonstrate that AFDEN performs better than variants lacking aspect-feature distillation (AFD), aspect-feature enhancement (AFE), and dual-feature extraction (DFE). The model's resilience to perturbations and noise in aspect words is further validated by the aspect robustness study conducted with the Aspect Robustness Test Set (ARTS).

3 PAPER TITLE : ITERATIVE NETWORK PRUNING WITH UNCERTAINTY REGULARIZATION FOR LIFELONG SENTIMENT CLASSIFICATION

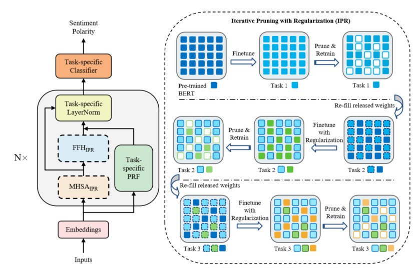

Figure 3: Illustration of the overall architecture based on the BERT model (left) and the processing flow of the iterative pruning with uncertainty regularization method (right).[3]

The paper proposes an iterative network pruning with uncertainty regularization method, called IPRLS, for lifelong sentiment classification using BERT as the base model. In order to let the model update old-task weights graciously, it includes three regularization strategies and incorporates uncertainty regularization based on the Bayesian online learning framework.The approach involves training an initial network for a new task, using weight-based pruning, and re-training the network after pruning to regain accuracy for the current task. Experimental results demonstrate that IPRLS outperforms strong baseline methods for lifelong sentiment classification across 16 popular datasets
Main Architecture: The main architecture involves leveraging network pruning and weight regularization to adapt a single BERT model to continuously arriving data from multiple domains. It incorporates uncertainty regularization based on the Bayesian online

learning framework and introduces three regularization schemes to guide the model to update old-task weights gracefully. The approach includes training an initial network for a new task, using weight-based pruning, and re-training the network after pruning to regain accuracy for the current task

Dataset: The paper evaluates the proposed method on 16 popular datasets for sentiment classification. However, specific details about the datasets are not provided in the summarized text.

Comparison with baseline: The proposed IPRLS approach and state-of-the-art (SOTA) methods are not directly compared in the paper. It does, however, contrast the IPRLS approach with a number other baseline techniques, such as TextCNN, BERT, PackNet, Piggyback, BiLSTM, and UCL. The experimental findings show that, for lifelong sentiment categorization across 16 widely used datasets, the suggested IPRLS technique performs better than these robust baseline methods. In order to compare the IPRLS method's performance against the baseline approaches, the accuracy of the lifetime sentiment classification is assessed.The experimental setup involves running all methods with the same task ordering during training, where the experimental data from 16 different domains arrive sequentially, and each dataset is considered as a separate task. The classification accuracy on the test set of each domain is reported after all 16 tasks are visited. This approach allows each model to keep learning a new task by using the weights learned from past tasks as initialization. The results of the experiments demonstrate that the IPRLS method achieves better performance compared to the baseline methods. Specifically, the IPRLS method exhibits better accuracy than PackNet and UCL on new tasks and old tasks, respectively. This indicates that the proposed IPRLS method is more resilient and keeps relatively stable accuracy throughout the learning stages, outperforming the baseline methods in terms of lifelong sentiment classification

Experimental results: The suggested IPRLS method's experimental findings show that it is superior to robust baseline techniques for lifelong sentiment classification. The approach successfully preserves prior knowledge while acquiring new tasks by achieving positive backward and forward knowledge transfer. The positive backward knowledge transfer is evidenced by the stability of accuracy over the entire learning process. Unlike traditional models such as BERT, which may suffer from catastrophic forgetting and exhibit fluctuations in performance, IPRLS maintains relatively stable accuracy throughout the learning stages. This stability indicates that the IPRLS method effectively retains knowledge from past tasks, preventing interference from new tasks and preserving the performance on previously learned tasks. Positive forward knowledge transmission is also demonstrated by the IPRLS approach. It efficiently applies the knowledge it has learned from previous tasks to help it learn new ones. The technique's capacity for positive forward transfer, which is essential for lifelong sentiment classification , is demonstrated by its ability to perform better on previous tasks while learning new ones. The paper performs an ablation research to examine the efficacy of each IPRLS component in addition to the experimental data.This study provides insights into the contributions of the iterative pruning strategy, uncertainty regularization, and the parallel residual function to the overall performance of the IPRLS method. By systematically evaluating the impact of each component, the study offers a comprehensive understanding of the method's effectiveness and sheds light on the key factors driving its superior performance in lifelong sentiment classification . The exploration of the effect of task ordering on lifelong sentiment classification further enriches the experimental findings. By analyzing the method's performance under different task orderings, the paper provides insights into the robustness and adaptability of the IPRLS method across various learning scenarios. This analysis enhances the understanding of the method's applicability in real-world settings where the order of incoming tasks may vary

4 CONCLUSION

This directed reading report concludes with a comprehensive examination of sentiment analysis and sentiment categorization based on a review of three seminal papers. The study began by explaining how the Tag-Assisted Transformer Encoder (TATE) network helps to alleviate the difficulties caused by missing modalities in multimodal sentiment analysis. The adaptive architecture of TATE, which includes a Common Space Projection Module and Tag Encoding Module, showed remarkable resilience while managing uncertainty resulting from missing or unclear data modalities.With the development of the Aspect Feature Distillation and Enhancement Network (AFDEN), the investigation further explored the terrain of aspect-based sentiment analysis. Through a nuanced dual-feature extraction process, aspect-feature distillation, and supervised contrastive learning, AFDEN emerged as a paradigm-shifting model, attaining state-of-the-art performance on benchmark datasets. Its ability to discern aspect-related and aspect-unrelated features showcased a sophisticated understanding of sentiment nuances within textual data.The Iterative Network Pruning with Uncertainty Regularization approach (IPRLS) for lifetime sentiment classification is the result of a trajectory of inquiry. IPRLS showed skill in updating old-task weights gracefully using uncertainty regularization and weight-based pruning, utilizing BERT as its basic model. The method demonstrated its effectiveness in tackling the problems associated with continuous learning in sentiment classification with its empirical superiority over strong baseline methods on 16 widely used datasets.

Overall, these publications provide major contributions to the developing field of sentiment analysis, providing nuanced solutions to numerous difficulties. The rigorous examination of methodology, designs, and experimental results offered in these works reveals the way forward for both scholars and practitioners.

A REFERENCES

[1] Tag-assisted Multimodal Sentiment Analysis under Uncertain Missing Modalities: arXiv:2204.13707 [cs.LG]

[2] Aspect Feature Distillation and Enhancement Network for Aspect-based Sentiment Analysis

[3] Iterative Network Pruning with Uncertainty Regularization for Lifelong Sentiment Classification: arXiv:2106.11197 [cs.CL]