David Danialy
Aidan Jones
Lawrence Nguyen
Uyen Nguyen
Christian Oh

Group 3 Significant Paper: Aspect-invariant Sentiment Features Learning: Adversarial Multi-task Learning for Aspect-based Sentiment Analysis
Bin Liang, Rongdi Yin, Lin Gui, Jiachen Du, Yulan He, Ruifeng Xu
CIKM 2020

Section Summaries

Section 1: Introduction

Aspect-based sentiment analysis (ABSA) basically finds descriptors in each string that provide sentiments for their corresponding aspects. For instance, the sentence "Alex was nice, but Mary was rude." has two aspects and two descriptors: Alex:Nice, Mary:Rude. There is already significant work done with "attention-based models" that can accomplish this task very well, but sometimes the descriptors don't have an aspect to correspond to. The core idea of the paper is to separate aspect-invariant sentiments from aspect-dependent ones, potentially improving the performance of ABSA models. The paper proposes to create an adversarial multi-task learning framework that distinguishes aspect-invariant and dependent sentiments, and extracts the sentiments from invariant ones. The dataset will have synthetic data that has had the aspects swapped, while the sentiments stayed the same, in order to balance the dataset, for which some aspects have abundant data instances while others have very few.

Section 2: Related Work

This section describes all of the papers that have extended research on ABSA technologies, and it states that almost all of them need aspect-dependent sentiments to function correctly. There is a need for a better network to utilize aspect-invariant sentiments correctly. The paper also describes the work that has gone into Adversarial multi-task learning, especially in the NLP field, and states that a "novel framework to extract aspect invariant sentiment features" may improve aspect-based sentiment classification.

Section 3: Methodology

This section describes how each module of the proposed architecture functions:

3.1: Synthetic Cross-Aspect Sample Generation

As stated in the introduction, the dataset will be balanced with the inclusion of synthetic data instances for aspects that are lacking in them, by replacing the aspect in a data instance with another aspect that shares similarity in how the sentiment describes it, or its "genre."

3.2: Adversarial Multi-Task Learning Framework

The adversarial network consists of one branch for aspect-based sentiment prediction and another for aspect-invariant features extraction, and will easily combine with existing ABSA models.

3.3: Multi-Task Learning for ABSA

There are four components that go into the multi-task learning framework:

1. Aspect Discrimination
    a. Sentiments can either be aspect-invariant or aspect-dependent, and this portion of the multi-task learning predicts the category of aspects that the sentiment might describe.
2. Synthetic Sample Discrimination
    a. This portion of multi-task learning determines whether a cross-aspect data instance is real or synthetic using cross-entropy loss minimization of predicted and true distributions.
3. Sentiment Discrimination
    a. Aspect-invariant sentiment features are extracted and fed into the pre-trained classifier, and noisy and aspect-dependent sentiment features are discriminated. This improves the performance of the classifier on aspect-invariant data.
4. Aspect-Based Sentiment Prediction
    a. Simply uses an existing, pre-trained ABSA model to feed the aspect-invariant sentiments that were extracted by the feature extractor into in order to improve their performance on aspects that have fewer data instances.

3.4: Adversarial Training

The adversarial relationship is between the feature extractor and the discriminator. The feature extractor will help the ABSA model while impeding the discriminator's ability to predict aspect labels. Aspect labels that cannot be detected by the discriminator combined with the features extracted by the feature extractor will be useful for improving the ABSA model's accuracy for aspect-based sentiment prediction

3.5: Three Schemes of our Framework

There are three possible configurations for the feature extractor and the classifier model:

1. Integration Structure
    a. Directly integrate the feature extractor with the ABSA model
2. Concatenation Structure
    a. A recurrent model Bi-LSTM is used for feature extraction which concatenates the extracted sentiment features to the hidden output of the ABSA model
3. Gate Fusion Structure
    a. Similar to Concatenation, but feature combination is done with a fusion gate that utilizes the sigmoid function.

Section 4: Experiments

4.1: Datasets and Experimental Setting

The dataset used is two benchmark datasets of the restaurants domain from Semeval 2015 Task 12 and Semeval 2016 Task 5. GloVe is used to initialize word embeddings for non-BERT models, and a pre-trained uncased BERT-base is used for all BERT models.

4.2: Comparison Models

This section simply lists all of the models to which the proposed model is compared.

4.3: Main Experiment Results

An ABSA model based on the Adversarial Multi-Task Learning Framework performed significantly better than the competitor models. All three of the possible schemes for the framework improve accuracy for vanilla BERT models. Existing models combined with the Adversarial Multi-Task Learning Framework perform better than the previous state-of-the-art. The main reason for this is the improved performance on aspect-invariant data instances because of the synthetic data samples.

4.4: Ablation Study

Just using synthetic samples in ABSA models isn't effective. Using the entire multi-task learning framework results in better scores, especially when using the synthetic samples. This is because discriminating noisy synthetic samples makes the better ones more effective. It is also significantly better to combine multi-task learning with adversarial training.

4.5: Impact of the Training Data Size

The efficacy of the synthetic training samples is shown by using varying amounts of annotated data compared to synthetic samples. When there is a significant lack of actual annotated data, the three framework schemes still perform acceptably. This means that the method of creating synthetic data is very useful for datasets with small amounts of annotated data. The accuracy of normal ABSA models have a negative correlation with the amount of synthetic data, while the three framework schemes proposed have a positive correlation.

4.6: Detailed Results for Aspect-Invariant Sentiment Extraction

90% of the synthetically created data instances that are discriminated wrongly have aspect-invariant sentiment features. Half of the data instances that are discriminated correctly have aspect-invariant sentiment features. This is why adding the synthetic data to a normal ABSA model gives a negative correlation with accuracy. It is fine on the proposed schemes, because they possess the adversarially trained discriminator to prevent the bad data from affecting the model training.

4.7: Visualizations and Qualitative Analysis

This section focuses on the proportions of discriminated synthetic data instances as described in the previous section, using visualizations to better display how the performance is distributed.

4.8: Analysis of Gated Fusion Structure (GFS)

The gated fusion structure scheme is the one that performed the best out of all three different schemes in all comparison experiments. This is because it better incorporates learning from aspect-invariant sentiment features.

4.9: Effect of Off-the-shelf Aspect-Invariant Data

With increasing availability for previously limited aspect-invariant data, the accuracy of baseline ABSA models will increase. This is because it will learn to better handle that data which was not available to train on in large numbers.

Section 5: Conclusion

The novel adversarial multi-task learning framework that was proposed in this paper is able to generate synthetic training data that will improve the performance of baseline ABSA models. It can easily integrate with the existing models to improve their performance to state-of-the-art levels.

List salient points from the paper

1. The method to create synthetic data using cross-aspect data is useful for training with very little annotated data.
2. The multi-task learning framework combined with the adversarial training provides the best possible results - better than the previous state-of-the-art.
3. A well-trained discriminator is a must for preventing the noisy synthetic data from affecting model training.
4. Current ABSA models don't handle aspect-invariant data well, which allows for the methods proposed in this paper to provide an increase in accuracy by combining those existing models with the multi-task learning framework.
5. Cross-aspect data is useful, as long as it isn't noise, for creating data instances where there previously were none or very few. This also allows for better handling of aspect-invariant data, when that cross-aspect data was created based on an aspect-invariant sentiment feature.

Appendix: Describe how we complete each of these

- Is it from a reputed journal or conference?
    - Yes, it is from CIKM 2020.
- Students summarized what they learned and the learnings are substantial
    - Evident from the summarizations of the sections, the level of understanding is sufficient.
- Are the contents relevant and significant to the course work?
    - Yes, this is a field of supervised learning and NLP that directly aligns with what we have learned in the class.
- Report captures salient points and is of good quality?
    - Yes, the salient points encompass the entire article's goal.
- Did the presentation make an impact? Did the students leave time and scope for Q&A and answer the questions confidently and correctly?
    - The presentation shows how the contents of the paper are the new state-of-the-art, making an impact, and it also leaves times for Q&A.
- Did the project build upon the significant paper?
    - Yes, our project is sentiment extraction using a BERT model, and we can integrate the adversarial multi-task learning network with it.
- Quality of slides - verbosity, readability, format, number, etc
    - The format is consistent, it describes everything in the fewest words possible, and the readability is good with sufficient font size and spacing. The number of slides is only as many as is needed to describe the contents of the paper.