Day 103: NLP Papers Summary – Utilizing BERT For Aspect-Based Sentiment Analysis Via Constructing Auxiliary Sentence

By Ryan 12th April 2020 No Comments

Objective and Contribution

Fine-tune pretrained BERT for Targeted Aspect-Based Sentiment Analysis (TABSA). TABSA is the task whereby you identify fine-grained opinion polarity towards a specific aspect associated with a given target. In ABSA, you don't have target-aspect pairs, just aspects. The contributions of this paper:

1. Proposed a new solution for TABSA by converting the task to a sentence-pair classification task via constructing auxiliary sentences
2. Fine-tuned pretrained BERT and achieve SOTA results on SentiHood and SemEval-2014 Task 4 dataset

Datasets

SENTIHOOD

- 5215 sentences, of which 3862 contain a single target and the remainder multiple targets
- Each sentence contains a list of target-aspect pairs with the sentiment polarity
- Given a sentence and the target, we need to
  - Detect mention of an aspect for the target
  - Determine polarity for the detected target-aspect pairs

Example:
LOCATION2 is central London so extremely expensive, LOCATION1 is often considered the coolest area of London.

Target	Aspect	Sentiment
LOC1	general	Positive
LOC1	price	None
LOC1	safety	None
LOC1	transit-location	None
LOC2	general	None
LOC2	price	Negative
LOC2	safety	None
LOC2	transit-location	Positive

Table 1: An example of SentiHood dataset.

SEMEVAL-2014 TASK 4

This is ABSA and not TABSA and so they don't have target-aspect pairs, just aspect. Allows the model to tackle subtask 3 (Aspect Detection) and subtask 4 (Aspect Polarity) at the same time.

Methodology

The paper uses 4 methods to construct auxiliary sentences to convert TABSA to a sentence pair classification task. Using (LOCATION1, safety) as a target-aspect pair:

- Sentences for QA-M

- The sentence we want to generate from target-aspect pair is a question and the f ^ ıat needs to be the same
- "What do you think of the safety of location – 1 ?"
- Sentences for NLI-M
  - The sentence created there is not a standard sentence but a simple pseudo-sentence that contains the target-aspect pair
  - "location – 1 – safety"
- Sentences for QA-B
  - Added label information and converted TABSA into a binary classification problem of yes or no to obtain the probability distribution
  - Each target-aspect pair will generate three sequences such as:
    - "the polarity of the aspect safety of location – 1 is positive"
    - "the polarity of the aspect safety of location – 1 is negative"
    - "the polarity of the aspect safety of location – 1 is none"
  - We use the probability value of yes as the matching score and take the class of sequence with the highest matching score for the predicted category
- Sentences for NLI-B
  - The difference between QA-B and NLI-B is that the auxiliary sentence changes from a question to a pseudo-sentence
    - "location – 1 – safety – positive"
    - "location – 1 – safety – negative"
    - "location – 1 – safety – none"

Methods	Output	Auxiliary Sentence
QA-M	S.P.	Question w/o S.P.
NLI-M	S.P.	Pseudo-sentence w/o S.P.
QA-B	{yes,no}	Question w/ S.P.
NLI-B	{yes,no}	Pseudo-sentence w/ S.P.

Table 2: The construction methods. Due to limited space, we use the following abbreviations: *S.P.* for *sentiment polarity*, *w/o* for *without*, and *w/* for *with*.

Experiments and Results

There are two experiment settings as we have two different datasets here.

SENTIHOOD EXPERIMENT SETUP

The evaluation of SentiHood only consider the 4 most frequently seen aspects in the dataset (general, price, transit-location, safety). The following are the chosen models for comparisons:

- Logistic Regression (LR)
  - LR with n-gram and pos-tag features
- LSTM-Final
  - biLSTM with final state as representation
- LSTM-Loc
  - biLSTM with the state associated with target position as representaion
- LSTM+TA+SA
  - biLSTM with complex target-level and sentence-level attention mechanism
- SenticLSTM
  - Upgraded version of LSTM+TA+SA, introducing external information from Sentic Net
- Dmu-Entnet
  - Bi-directional EntNet with external memory chains with a delayed memory update mechanism to track entities

SENTIHOOD RESULTS

- BERT-single outperformed Dmu-Entnet in aspect detection but scored a lower accuracy in sentiment classification accuracy

- BERT-pair outperformed other models on aspect detection and sentiment classificatio y a substantial margin. BERT-pair-QA models tend to perform better on sentiment analysis whereas BERT-pair-NLI models tend to perform better on aspect detection

SEMEVAL-2014 TASK 4 EXPERIMENT SETUP

BERT-pair models are compared against the best performing systems, namely, XRCE, NRC-Canada, and ATAE-LSTM.

SEMEVAL-2014 TASK 4 RESULTS

- BERT-single model was enough to achieved better results on the two subtasks
- BERT-pair models achieved a further improvements over BERT-single model

Conclusion and Future Work

WHY IS THE EXPERIMENTAL RESULT OF BERT-PAIR SO MUCH BETTER?

- The conversion of target-aspect pairs to auxiliary sentences is equivalent to expandi he corpus, therefore, more data for our models
- BERT seems to perform well in QA and NLI tasks indicating an advantage in dealing with sentence pair classification. This might be attributed to unsupervised masked language model and the next sentence prediction tasks
- Directly fine-tuning the pretrained BERT on TABSA does not yield good results. By transforming TABSA into a sentence pair classification task, the context is now similar to QA and NLI and so the advantage of pretrained BERT can be utilised

The conversion from single sentence classification to sentence pair classification task has yielded strong results. Future work could apply this conversion method to other similar tasks.

Source: https://arxiv.org/pdf/1903.09588.pdf

Ryan
Data Scientist

Next Post

Day 104: NLP Papers Summary - SentiHood: Targeted Aspect Based Sentiment Analysis Dataset

Previous Post

Day 102: NLP Papers Summary - Implicit and Explicit Aspect Extraction in Financial Microblogs

for Urban
Neighbourhoods

Data Science Natural Language Processing NLP Papers Summary

Day 365: NLP Papers Summary – A Survey on Knowledge Graph Embedding

Ryan
30th December 2020

Data Science Implementation Natural Language Processing

Day 364: Ryan's PhD Journey – OpenKE-PyTorch Library Analysis + code snippets for 11 KE models

Ryan
29th December 2020

Data Science Ryan's PhD Journey

Day 363: Ryan’s PhD Journey – Literature Review – Knowledge Acquisition – 1st Passes XIV

Ryan
28th December 2020