

Applied Natural Language Processing

Semester 1, 2022 Assignment 3: Building an aspect-based sentiment analysis algorithm based on syntactic parsing

DUE: 14 June. 2022, Tuesday 11:55 PM

Submission

Instructions and submission guidelines:

- You must sign an assessment declaration coversheet to submit with your assignment.
- Submit your assignment via the Canvas MyUni.

Task

You are required to write code for building an aspect-based sentiment analysis (ABSA) system based on syntactic parsing. The aspect-based sentiment analysis try to analyse the sentiment, e.g., positive, negative and neutral, toward a given aspect (aspect term). The following shows an example of ABSA:

For example:

"I loved their **fajitas**" → {fajitas: *positive*}
"I hated their **fajitas**, but their **salads** were great" → {fajitas: *negative*, salads: *positive*}
"The **fajitas** are their first plate" → {fajitas: *neutral*}

In this assignment, you will work on a dataset **with the aspect term given** (such as fajita in the above example), your task is to design a system to predict the three possible sentiment polarities: positive, negative and neutral with respect to the aspect term.

You are required to define rules based on syntactic parsing results (dependency parsing or constituency parsing or both) to achieve the sentiment prediction.

The following shows an example rule

Rule:

If the aspect term's child is "good or beautiful or delicious or yummy" with dependency type "amod" then the sentiment towards the aspect term is positive.

For a given rule, you can then analyse its precision and recall based on the ground-truth labels given in the dataset. For examples, if the rule is to detect positive sentiment, then the precision is calculated via

Precision = The number of detected true positive sentences (i.e., detected positive and ground-truth is also positive) /The number of sentences that are detected positive

Recall = The number of detected true positive sentences/ The total number of sentences with positive sentiment

You are required to identify multiple rules (more than 2 for each polarity) for aspect-based sentiment analysis. You need to evaluate the precision and recall of each rule, and submit your discovery as a report with accompany code (you need to submit both the report and code).

The detailed rubric is as follows

Rubric

Aspects	0-49	50-79	80-100
The correctness of implementation 50%	Incomplete Simple implementations with errors or code does not match the results in the report		Correct implementations.
The diversity of the identified rules 25%	Only consider few simple rules	Use more than one rules per polarity	Use diverse rules. Each rule is based on quite different patterns
Report presentation 25%	Only show the required numbers		Also give examples about successful or failure cases as the analysis

Data

You will use Restaurant review dataset which is provided at MyUni

The xml files (please Google how to load xml file with python) which have multiple fields, you only need to consider the following fields

Sentence, Aspect Term and Polarity

Please note that the span (based on characters) of the aspect term is also given in the dataset for your convenience.