

Annotation Guidelines for Dataset DAORA

1 Introduction

This document contains annotation guidelines for *Dataset for Aspect, and Opinion Review Analysis - DAORA* and *Multi-level Dataset for Aspect, and Opinion Review Analysis - ML-DAORA*. This annotation guidelines were inspired by guidelines created for SemEval tasks for aspect-based sentiment analysis (ABSA) (Pontiki et al., 2014; Pontiki et al., 2015; Pontiki et al., 2016). The main goal of this task is to identify opinions expressed in hotel reviews towards entities and their attributes. In particular, given a hotel review, the task of the annotator is to identify the following types of information:

- **Aspect Term (Opinion Target) Expression (ATE)** is an explicit reference (mention) to the reviewed entity. This reference is identified by its starting and inside label.
- **Opinion Term Expression (OTE)** is an explicit reference (mention) of subjective expression towards reviewed aspect term. This reference is identified by its starting and inside label.
- **Aspect Term Polarity (ATC)** - Each aspect term has to be assigned one of the following classes based on the sentiment that is expressed in the sentence about it:
 - no sentiment (0) - only for non-aspect terms
 - negative (1)
 - neutral (2)
 - positive (3)
 - conflict (4)

2 Aspect Term Expression

An aspect term expression (ATE) is an explicit reference (mention) to an entity that is evaluated. This mention can be a named entity, a common noun, or a multi-word term. In the case of multi-word terms, the maximal phrase should be tagged.

2.1 What should be annotated as aspect term:

1. Nominal phrases explicitly mentioning aspects.
 - (a) You will get nice , clean, comfortable modern accommodations in a fantastic location.
2. References to the target entity (the hotel the review is about including its restaurant etc.), except mentions of other entities (e.g. sightseeing points, other hotels, or restaurants which the review is not about).
 - (a) This **hotel** is just excellent !!
 - (b) The **hotel** is a bit outdated...
 - (c) Nice hotel , fantastic place, expansive parking.
 - (d) There are plenty of restaurants nearby including the Hard Rock Cafe just across the canal.
 - (e) Marriot **Hotel** is... and Hyatt Hotel is...¹
3. If an aspect term appears in a sentence more than once, all of its occurrences in the sentence should be annotated.
 - (a) The **room** we had was very spacious, warm and quiet, even though our **room** was right behind the elevator, and conveniently next to an ice machine.

¹In case the review is about Marriot

2.2 What should not be annotated as aspect term:

1. Terms that are often used as aspect terms (e.g., “room”) may not always be aspect terms.
 - (a) there was not enough room for bags
2. Subjectivity indicators - should not be tagged as part of aspect term but as an opinion term.
 - (a) Overpriced **breakfast** .
 - (b) Good **spreads**, great **beverage selections** and bagels *really tasty*
 - (c) Both the fresh mozzarella slices and the Plain Cheese slice are phenomenal²
3. The name of the hotel or any other entity.
 - (a) Hilton Berlin - It 's fine . . .
4. Pronouns (e.g., “it”, “they”, “this”) even if they refer to an aspect.
 - (a) We went for **breakfast** one day. It was fresh and delicious.
5. Implicit aspect terms - aspect terms that are not explicitly mentioned, but can be inferred from adjectives or other expressions.
 - (a) Overall though, it was good.
6. Determiners (e.g., “a”, “the”, “some”, “many”, “all”) should not be included in aspect terms.
 - (a) Some **rooms** are bigger and nicer than others.
7. Verbs or verbals (words formed from a verb, but functioning as a different part of speech). Those forms should not be tagged as aspect term (a) but as opinion term (b). (a) is also an example of an implicit aspect term of an entity without explicit mention in the sentence.
 - (a) Fresh, delicious, and reasonably priced
 - (b) the **minibar** was very reasonably priced
8. Enumerations. “fresh meats, cheese, pastries, yoghurt [...]” should not be tagged.
 - (a) We paid for our **breakfast** as part of our package and the **choice** was very impressive, fresh meats, cheese, pastries, yoghurt, fruit juice and pancakes (served with a choice of sauces, my favourite was chocolate ! ! !)
9. When an aspect term is part of factual information and has no sentiment.
 - (a) Our room was in the newer tower on the 18th floor.
10. Expectations - expectations of a reviewer which have not been satisfied.
 - (a) I expect to have my room made up.
11. Suggestions.
 - (a) I would suggest a microwave in the room.

3 Opinion Term Expression

An opinion term expression (OTE) is an explicit reference (mention) of a subjective expression towards the reviewed aspect term. This mention is in most cases an adjective. Unless more adjectives are separated by commas or conjunctions, all adjectives should be annotated as one opinion term expression.

²Note that some terms can be used both as parts of aspect terms as well as subjectivity indicators (e.g., “fresh” in the following sentences).

3.1 What should be annotated as opinion term:

1. Subjectivity indicators - should be tagged as part of opinion term.
(a) *Overpriced* **breakfast**.
2. Adjectives separated by commas or conjunction are annotated as separate opinion terms.
(a) You will get *nice, clean, comfortable modern* **accommodations**...

3.2 What should not be annotated as opinion term:

1. Sentences in which an aspect term has no explicit reference to the opinion term.
(a) ... even though our **room** was right behind the elevator, and conveniently next to an ice machine
2. Subjectivity indicators which cannot be assigned to specific aspect terms unless they could be assigned to other aspect terms in a connected sentence.
(a) Fresh, delicious, and reasonably priced.
(b) We went for **breakfast** one day. It was *fresh* and *delicious*.

4 Aspect Term Polarity

In Aspect Term Polarity, we consider annotations as follows:

- No polarity - with no polarity are annotated all tokens expect aspect terms
- Positive and negative polarity - aspect terms are annotated with the polarity of the associated opinion term expressions if all have the same polarity.
 - The concierge was very helpful. (positive)
 - Service was mediocre at best. (negative)
- Conflicting polarity: aspect terms are annotated with conflicting polarity if they are connected to both, positive and negative sentiment OTEs.
 - The **hotel staff** was *friendly* but *not very helpful*. (conflict)
- Neutral polarity - when expression like “moderate”, “in line”, “nothing out of the ordinary”, “not an issue” etc. are used. ATEs with connections to neutral and either positive or negative OTEs are annotated in either positive or negative.
 - The **room** itself was *just ordinary*. (neutral)
 - *good* **hotel** but *not perfect* . (neutral)

References

- Maria Pontiki, Dimitris Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. SemEval-2014 task 4: Aspect based sentiment analysis. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 27–35, Dublin, Ireland, August. Association for Computational Linguistics.
- Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos. 2015. SemEval-2015 task 12: Aspect based sentiment analysis. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 486–495, Denver, Colorado, June. Association for Computational Linguistics.
- Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammad AL-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphée De Clercq, Véronique Hoste, Marianna Apidianaki, Xavier Tannier, Natalia Loukachevitch, Evgeniy Kotelnikov, Nuria Bel, Salud María Jiménez-Zafra, and Gülşen Eryiğit. 2016. SemEval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 19–30, San Diego, California, June. Association for Computational Linguistics.