SemEval 2015 Task 12

Aspect-Based Sentiment Analysis (ABSA)

Submission guidelines

The ABSA 2015 task will run in two phases. In the first phase (phase A), participants will be given two test sets, one for the laptops and one for the restaurants domain, both in the same xml format. The test sets will contain sentence spitted review texts. The restaurant sentences should be annotated with the identified <target, aspect category> tuples, whilst the laptop sentences should be annotated only with the identified aspect categories. In the second phase (phase B), participants will be given the gold (correct) annotations for the above two datasets, as well as for a hidden domain (subtask 2) and they will have to return the corresponding polarities (positive, negative, neutral). The evaluation measures for all slots are described in the ABSA15BaseEvalValid.pdf of the *Baselines & Evaluation package*[1].

In the <Reviews> tag of the submitted xml files you should add the appropriate attribute(s) that indicate whether you participate in a slot and whether the submission is constrained or unconstrained. For example, in the case of the restaurants domain where two slots are required in phase A, you have to add slot1, slot2 and slot12 attributes for Slot1, Slot2 and <Slot1, Slot2> respectively. A "" value means that the submission will not be evaluated in the corresponding slot(s) while "c" and "u" indicate that the submission for this slot is will be evaluated as constrained or unconstrained, respectively. Each team may submit up to two runs per slot of each domain, one constrained and one unconstrained[2]. In both phases (A and B), the xml file(s) that your system will generate should have the same format with the training files. A script that validates that a system's output (xml) is in the correct format is provided in the *Baselines & Evaluation package*. All your xml files for each phase should be uploaded as a single zip. The uploaded zip file should be named **TID_PH.zip**, where **TID** is the id of your team and **PH** is the phase of the evaluation ("A" or "B"). Each submitted file name should have the following format **TID_PH_DOM_SUBN.xml**, where **DOM** is the domain of the submission ("Lap" for Laptops "Res" for Restaurants and "Hid" for the hidden domain). **SUBN** indicates the number of submission. For example, a team with id "tm1" that participates in phases A and B for all domains with two runs it should upload the following 10 xml files archived in two zip files:

tm1_A.zip	tm1_A_Lap_1.xml, tm1_A_Lap _2.xml, tm1_A_Res_1.xml, tm1_A_Res _2.xml
tm1_B.zip	tm1_B_Lap_1.xml, tm1_B_Lap_2.xml, tm1_B_Res_1.xml, tm1_B_Res_2.xml, tm1_B_Hid_1.xml, tm1_B_Hid _2.xml

Participants are free to decide the domain(s), subtask(s), and slot(s) they wish to participate in. In the case of restaurants, where there are two slots for Phase A you can:

- Submit results only for Slot1. Then for the sentence "*The prices were CHEAP compared to the quality of service and food*" the following annotations should be returned.

```
<Opinion target="" category="RESTAURANT#PRICES" />
<Opinion target="" category="SERVICE#GENERAL" />
<Opinion target="" category="FOOD#QUALITY" />
```

If it is an unconstrained submission, the attributes in <Reviews> tag of the xml file should be filled as follows: <Reviews slot1="u" slot2="" slot12="">

- Submit results only for Slot2. Then for the previous sentence the following annotations should be returned.

```
<Opinion target="service" category="" from="49" to="56"/>
<Opinion target="food" category="" from="61" to="65"/>
```

If it is a constrained submission, the attributes in the <Reviews> tag of the xml file should be filled as follows: <Reviews slot1="" slot2="c" slot12="">

- Submit results for Slot1 and Slot 2 but not <Slot1, Slot2> tuples. Then for the previous sentence the following annotations should be returned.

```
<Opinion target="" category="RESTAURANT#PRICES" />
<Opinion target="" category="SERVICE#GENERAL" />
<Opinion target="" category="FOOD#QUALITY" />
<Opinion target="service" category="" from="49" to="56"/>
<Opinion target="food" category="" from="61" to="65"/>
```

If slot2 is constrained and slot1 is unconstrained, the attributes in the <Reviews> tag should be filled as follows: <Reviews slot1="u" slot2="c" slot12="">

- Submit results for Slot1, Slot 2 and <Slot1, Slot2> tuples. Then for the previous sentence the following annotations should be returned.

```
<Opinion target="NULL" category="RESTAURANT#PRICES" polarity="positive" from="0" to="0"/>
<Opinion target="service" category="SERVICE#GENERAL" polarity="positive" from="49" to="56"/>
<Opinion target="food" category="FOOD#QUALITY" polarity="positive" from="61" to="65"/>
```

If slot1 is constrained and slot2 is unconstrained, the attributes in the <Reviews> tag of the xml file should be filled should be as follows: <Reviews slot1="u" slot2="c" slot12="u">

In the case of laptops for Phase A you will have only to add slot1 attribute. Similarly in phase B for all domains (laptops, restaurant, hidden) you will only have to add slot3 attribute.

The submission process will be as follows: Initially you will be sent a link to a page, where you will fill the required information (e.g. team ID, affiliation) and register for the task. After that you will receive an e-mail that confirms your registration. The same e-mail will include the URLs for the test data, it will describe how you will upload the zip file with your xml files and also it will explain how you will submit a brief description of your system. Test data for **phase A** will be available at **15:00 PM (GMT) on December 16** and participants will have to submit their system's output until **15:00 PM (GMT) on December 18**. Test data for **phase B** will be released at **09:00 AM (GMT) on December 19** and participants will have to submit their system's output until **21:00 PM (GMT) on December 22**