Aspect Based Sentiment Analysis

Ankit Singh 12128
Enayat Ullah 12407

March 2015

1   Motivation

Sentiment analysis aims to determine the evaluation of an author with respect
to a particular topic and detecting the overall contextual polarity of a document.
Sentiment analysis is increasingly viewed as a vital task both from an academic
and a commercial standpoint. The majority of current approaches, however,
attempt to detect the overall polarity of a sentence, paragraph, or text span,
irrespective of the entities mentioned (e.g., laptops) and their aspects (e.g.,
battery, screen).
The project is directed towards aspect-based sentiment analysis, where the goal
is to identify the aspects of given target entities and the sentiment expressed
for each aspect.

2   Problem Statement:

The task is focussed on two domains: restaurants and laptops. For each of these
domains, we have a predefined set of entity types(E) and attributes(A). For lap-
tops, we have: 22 entity types and 9 attribute labels, and 6 entity types and 5
attribute labels for restaurants. The E#A pair defines an aspect (category).

The following tasks would be addressed in this project:

2.1   Task 1: In domain ABSA

Sub-task 1:
Identification of every entity E and attribute A pair E#A towards which an
opinion is expressed in the given text. E and A are chosen from predefined in-
ventories of Entity types (e.g. laptop, keyboard, operating system, restaurant,
food, drinks) and Attribute labels (e.g. performance, design, price, quality) per
domain.

Sub-task 2:

Identification of Opinion Target Expression (OTE). An opinion target expression (OTE) is an explicit reference (mention) to the reviewed entity E of the EA pair. The OTE slot takes the value "NULL", when there is no explicit mention of the opinion entity. Each identified E#A pair of the given text is then assigned a polarity, from a set P = positive, negative, neutral

2.2   Task 2: Out of domain ABSA

(If time permits)
Domain adaptation considers the setting in which the training and testing data are sampled from different distributions.The implementation will be tested on a previously unseen domain for which no training data is made available.

3   Methodology:

3.1   Task 1:

Sub-task 1:Aspect term detection
Using Conditional Random Fields (CRF) with features extracted using named entity recognition, POS tagging, parsing, and semantic analysis.[2],[3]

Sub-task 2: Aspect polarity detection
Using an SVM classifier with features mainly based on n-grams, parse trees, and several out-of-domain, publicly available sentiment lexica (e.g. Wordnet, SentiWordnet).[4]

3.2   Task 2:

(If time permits)
A deep learning approach can be used to address the problem of domain adaptation for sentiment classifiers. A higher-level feature extraction is learnt in an unsupervised fashion from the text reviews of all the available domains using a Stacked Denoising Autoencoder(SDA). We then use an SVM to adjudge the polarity of the sentence.[5]

4   Dataset:

Two datasets of 550 reviews of laptops and restaurants annotated with opinion tuples are provided by SemEval in XML format. All possible E#A pairs are explicitly mentioned for the laptop and restaurant domain.[1]

5   References:

[1].SemEval-2015 Task 12: Aspect Based Sentiment Analysis
`http://alt.qcri.org/semeval2015/task12/`

[2].Identifying Sources of Opinions with Conditional Random Fields and Extraction Patterns, Yejin Choi and Claire Cardi, Ellen Riloff and Siddharth Patwardhan
`https://www.cs.cornell.edu/home/cardie/papers/hlt-emnlp05-yejin.pdf`

[3].Dependency Tree-based Sentiment Classification using CRFs with Hidden Variables, Tetsuji Nakagawa,Kentaro Inui and Sadao Kurohashi
`http://www.aclweb.org/anthology/N10-1120`

[4]. Sentiment analysis using support vector machines with diverse information sources, Tony Mullen and Nigel Collier
`http://www.aclweb.org/old_anthology/W/W04/W04-3253.pdf`

[5].Domain Adaptation for Large-Scale Sentiment Classification: A Deep Learning Approach, Xavier Glorot, Antoine Bordes, Yoshua Bengio
`http://www.icml-2011.org/papers/342_icmlpaper.pdf`