

Feature based opinion mining

One of the basic tasks in opinion mining or sentiment mining is classifying the polarity of a given text or feature/aspect level to find out whether it is positive, negative or neutral. Different methodologies are used for this purpose. Some expert analysts used the scaling system to associate numbers with appropriate sentiments that a word is depicting. Research has also shown that subjectivity or objectivity identification can also achieve the purpose. However the most fine grained analysis model would be the feature or aspect based sentiment mining method for this purpose.

The basic idea of feature based sentiment mining is to determine the sentiments or opinions that are expressed on different features or aspects of entities. When text is classified at document level or sentence level it might not tell what the opinion holder likes or dislikes. If a document is positive on an object it clearly does not mean that the opinion holder will hold positive opinions about all the aspects or features of the object. Similarly if a document is negative it does not mean that the opinion holder will dislike everything about the object described.

Tasks in Feature Based Opinion Mining

When evaluating documents like product reviews mostly the opinion holders write both positive and negative reviews about a product or service. The general sentiment may be evaluated with the mood of the document to be positive or negative. However in order to obtain such detailed aspects in a text feature level evaluation is required. There are generally three key types of tasks involved in feature based opinion mining:

1. **Identifying the Object Features:** If we are analyzing a review of a product like digital camera and it says that 'the picture quality of this camera is good' the object feature here will be picture quality. This is the technique involved here to identify nouns and noun phrases as features. Here information extraction techniques can be valuable like Conditional Random Fields or Hidden Markov Models.
2. **Determining the Opinion Orientation:** The second task is to determine whether the feature that is outlined as positive, negative or neutral. This can be slightly complicated in some sentences but achievable generally. For example, the opinion on the feature picture quality is positive. Many approaches can be used in this regard. A commonly known approach is the lexicon based approach for determining the orientation of an opinion or feature in sentences. Another approach like relaxation labeling based approach can also be used.
3. **Grouping Synonyms:** After determining the orientation, the third task is to group the synonyms. The object features can be expressed with different words and phrases.

The overall method of feature based sentiment mining is simple to understand, however during the process many questions may arise like how the set of features can be obtained and how are they linguistically expressed. Their relation to each other is also important.

Knowledge representation models can be used to organize features in a better way. These models can be helpful in structuring features, extracting the explicit and implicit features and producing a discourse based summary as well. A well known model in this aspect is the rule mining algorithm for discovering product features through nouns and noun phrases.

Much research has been done to find methods of improving feature extraction to a refined level. The methods of ontology and lexicon have been useful in this regard but still work is being done to improve the overall effectiveness of feature based opinion mining.