

Chapter 18

Ontology-Based Opinion Mining

Rajendra Akerkar

Western Norway Research Institute, Norway

Terje Aaberge

Western Norway Research Institute, Norway

ABSTRACT

In this chapter, the authors discuss an ontology-based approach to opinion mining exploiting the possibility to represent commonly shared meaning of linguistic relations by ontologies. The ontology definitions are used as a standard to which sentences extracted from texts are compared. Unlike conventional text mining, which is based on objective topics aiming to discover common patterns of user opinions from their textual statements automatically or semi-automatically, it will extract opinion from subjective locations.

INTRODUCTION

The rapid growth in Internet applications has led to an enormous amount of personal reviews on the Web covering many subject areas. These reviews appear in different forms like BBS, blogs, Wiki or forum websites and present information valuable to both the general public searching providers of services and providers trying to adapt to public demands. An intrinsic problem of the overwhelming information, however, is information overloading as users are simply unable to access and read

all the available information. Query functions in search engines like Yahoo and Google help users to find reviews but the identification of relevant information from returned pages might demand more effort than worthwhile.

The totality of reviews on an offer may be assumed to represent public opinion concerning its quality. The importance of knowledge of public opinion makes methods of extraction, i.e. opinion mining and sentiment analysis, enjoy a burst research activity. Opinion Mining is the field that deals with the mining of subjective statements from

texts, the identification of opinions, the estimation of opinion orientation and the extraction of arguments that relate to opinions. Mining opinions in online discussions requires an appropriate representation. It is distinctively different from traditional text mining in that the latter is based on declarative descriptions while the opinion mining must take into account the additional level of description of subjective experience. Specifically, traditional text mining focuses on specific topics (e.g., business, travel) as well as topic shifts in text whereas opinion mining is much more difficult than those for topic mining. This is partially attributed to the fact that topics are represented explicitly with keywords while opinions are expressed with subtlety. Opinion mining requires deeper understanding of language characteristic and textual context (Pang, et al., 2002).

In the past few years, many researchers studied the problem. Most of the existing works are based on product reviews because a review usually focuses on a specific product and contains little irrelevant information. The main tasks are to discover properties that have been commented on and to decide whether the opinions are positive or negative. In this chapter, we discuss the property identification problem by employing ontology structure. Ontology allows us to analyse a review at a finer granularity based on shared meaning.

An ontology provides knowledge about domains that is understandable by developers and can be handled by computers. It is prevalent in annotating documents with metadata, improving the performance of information retrieval and reasoning, and making data interoperable between different apps (Akerkar, 2009). Ontology-based text analysis seems promising for opinion mining.

BACKGROUND

There are two fundamental research efforts in opinion mining area that have been given attention, namely sentiment classification and

property-level opinion mining. Traditional document classification methods are often employed in sentiment classification; the entire document is classified as positive, negative, or neutral. However, these methods must be considered too vague. In most cases, both positive and negative opinions can appear in the same document, for example, “Overall, the accommodation is super, the glacier-hike is thrilling, but the food is tasteless.” There are some works focused on sentiment classification. They are based on manually or half-manually constructed a priori knowledge dictionary which contains polarity words. Some classic machine learning methods (Naive Bayes, Maximum Entropy, and SVM) have been tested by Pang et al. (2002) who point out that machine learning methods are more effective than manually labelling. Hearst (1992) proposed a model-based approach trying to analyze the structure of the document more deeply, but the method is so far not supported by experimental data.

In the property level opinion mining critic extraction-related sub-tasks were identification of properties and extraction of opinions associated with these properties. Existing works on identifying object properties discussed in reviews often rely on the simple linguistic heuristic that properties are represented by nouns or noun phrases. Popescu and Etzioni (2005) consider product properties to be predicates forming relationships with the product (for example, for a tour planning, transport is one of its properties) and try to identify the properties connected with the product name. Their approach does not involve opinion mining but simply focuses more on the task of identifying different types of properties. Hu and Liu (2004a) propose the idea of opinion mining and summarization. They use a lexicon-based method to determine whether the opinion expressed on a product property is positive or negative. In a later paper Hu and Liu (2004b) follows the intuition that frequent nouns or noun phrases are likely to be properties. They identify frequent properties through association mining

10 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/ontology-based-opinion-mining/71864

Related Content

Semantics for the Semantic Web: The Implicit, the Formal and the Powerful

Amit Sheth, Cartic Ramakrishnan and Christopher Thomas (2005). *International Journal on Semantic Web and Information Systems* (pp. 1-18).

www.irma-international.org/article/semantics-semantic-web/2802

Wiki Semantics via Wiki Templating

Angelo Di Iorio, Fabio Vitali and Stefano Zacchiroli (2010). *Handbook of Research on Web 2.0, 3.0, and X.0: Technologies, Business, and Social Applications* (pp. 329-348).

www.irma-international.org/chapter/wiki-semantics-via-wiki-templating/39179

Semantics for the Semantic Web: The Implicit, the Formal, and the Powerful

Amit Sheth, Cartic Ramakrishnan and Christopher Thomas (2007). *Semantic Web-Based Information Systems: State-of-the-Art Applications* (pp. 1-21).

www.irma-international.org/chapter/semantics-semantic-web/28907

Ontologies with Semantic Web/Grid in Data Integration for OLAP

Tapio Niemi, Santtu Toivonen, Marko Niinimäki and Jyrki Nummenmaa (2007). *International Journal on Semantic Web and Information Systems* (pp. 25-49).

www.irma-international.org/article/ontologies-semantic-web-grid-data/2841

Supporting Conceptual Model Analysis Using Semantic Standardization and Structural Pattern Matching

Patrick Delfmann, Sebastian Herwig, Lukasz Lis and Jörg Becker (2012). *Semantic Technologies for Business and Information Systems Engineering: Concepts and Applications* (pp. 125-149).

www.irma-international.org/chapter/supporting-conceptual-model-analysis-using/60059