Aspect-oriented sentiment analysis of hotel reviews

Nur-Sultan,L.N. Gumilyov Eurasian National University

CREATED BY:

ORALBEKOVA INGKAR

YERGESH BANU

Contents

Introduction

▶ Currently, customers do not need to call the hotel staff to check the availability of rooms, prices and connecting transport. Instead, they can simply browse the hotel's website and other commercial platforms (for example, Booking.com).They can book a room, pay a bill and share reviews about services, all this online with a few clicks.

▶ Basically, these sites rank hotels according to the "Ratings" that some of the previous travelers provide based on reviews. Despite the fact that some websites may provide ratings based on some aspect, but still research shows us that interaction with service providers and other customers is necessary to properly meet customer needs[1].

▶ The reliability of these ratings is also questionable [2]. This requires user reviews mentioned on the website that give a good idea of what other customers have experienced while visiting this particular hotel. Customers are more sensitive to the personal information they find in reviews, and most often use it as a basis for making decisions.

- Most customers are likely to read the first fifty or the most popular reviews, which may lack some of the features described by others. This study analyzes data from hotel reviews and uses various natural language processing methods to identify some important information that is apparently not visible to the viewer.

- The discussion of a small background is followed by an in-depth study of brief concepts of text analysis. Then some related work that has been done in this area is discussed, based on the proposed approach to solving the problem. This is followed by a discussion of the experiments conducted and the results obtained, summarizing the results regarding the future scope of improvement.

Text Analytics

▶ Due to unpredictable size of review sets and customer generated content, various text analysis techniques like sentiment analysis, opinion mining, topic modeling, classification, etc. play an important role.

▶ Classification is an important step and is a supervised method and is used to classify the text into various classes. In our study, we use it to classify the sentences of the text into predefined aspects that we already know are present in hotel reviews.

▶ Topic Modeling is a type of machine learning technique that can be used as a statistical model to discover the abstract topics in a collection of documents. It is a valuable tool to identify hidden semantic structures in the textual data and performs this function over our dataset[6].

▶ Sentiment Analysis can be used easily for extracting opinions from the data about a certain aspect[7]. It is particularly useful for unstructured human authored documents and is a very important factor in business intelligence. It has become the central part of Information Retrieval process. The approaches towards short text summarization have improved sentiment analysis techniques [8].

Related Work

- OpeNR[10] is an NLP platform applied to the hospitality domain to automatically process customer-generated reviews and extract important information from it. It consists of a set of OpenSource and free NLP tools to analyze text based on a modular architecture to simplify its modification and extension. It basically takes in social media generated content to perform text analysis on reviews[11].

- Then there is system which is a part of the BESAHOT[16] project[15]. It is meant to be targeted at hotel customers who wish to know the information and actual overviews and summaries of the textual content about their hotels on the web. It handles only German reviews from German websites as of the time of this writing. It is based on the GWT framework and is an interactive web application. The core system handles data acquisition, analysis and storage while the user interface provides various types of summaries of the analyzed data.

Aspect Oriented Sentiment Analysis

▶ Reading all reviews is time consuming, therefore the need for automation

▶ Sentiment mining is defined as the automatic assessment of the sentiment expressed in text (in our case by consumers in hotel reviews)

▶ Several granularities of sentiment mining:

– Review-level

– Sentence-level

– Aspect-level (hotel aspects are sometimes referred to as hotel features): Aspect-Based Sentiment Analysis (ABSA):

• Review-level

• Sentence-level
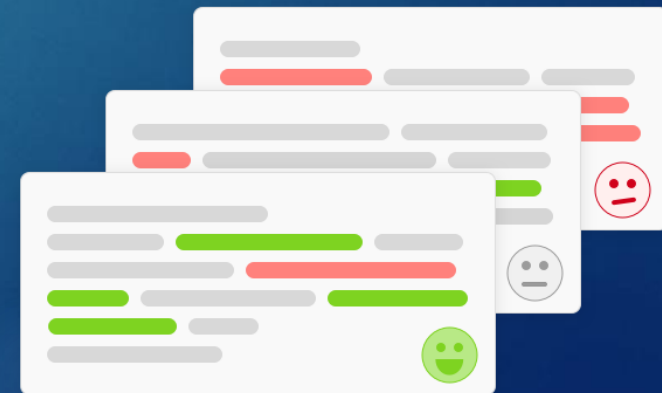
Aspect Oriented Sentiment Analysis

▶ Aspect-based sentiment analysis (ABSA) is a text analysis technique that categorizes data by aspect and identifies the sentiment attributed to each one.

▶ In the textual data, the entity may be described using different sets of words aka features detailing the entities; these features are called 'aspects' of the corresponding entity. ABSA strives to find the relevant aspects detailing the entity along with their upholding words and further determines the polarity or sentiment of those aspects.

Example:

Negative

Sentence: **Бөлме керемет** болды, бірақ **қызметкерлер жағымсыз** болды

Positive

Sentence: The **room** was **great**, but the **staff** was **unfriendly**

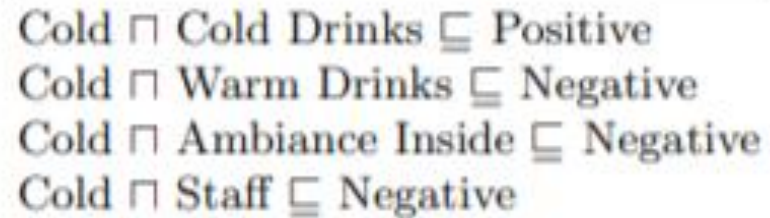| Aspect category	Aspect term	Sentiment
Room	Great	Positive
Service	Unfriendly	Negative

Data

- ▶ • Each review-aspect pair is annotated with sentiment: positive, negative, or neutral

- ▶ • A sentence or review can contain multiple aspects

- ▶ • Task: detect the aspect-based sentiment at review-level

Hotel aspects

Aspect	Set of values
Room	{-2; -1; 0; 1; 2}
Location	{-2; -1; 0; 1; 2}
Service	{-2; -1; 0; 1; 2}
Food	{-2; -1; 0; 1; 2}

Ontology

▶ The ontology can be divided into two parts, corresponding with two top-level classes: Mention and Sentiment. The class Mention has two subclasses, each with its own subclasses. The first subclass of Mention is Entity, with its domain specific subclasses Ambiance, Experience, Location, Person, Price, Restaurant, Service, Style Options, and Sustenance. The domain-specific subclasses are annotated with the corresponding aspect categories, which is mostly a one-to-one mapping. Furthermore, within these classes, we group certain concepts together with the intention to improve precision. For example, Warm Drink and Cold Drink as subclasses of Drink.

▶ The second subclass of Mention is Property. Its subclasses represent different properties of entities. These subclasses are constructed according to the Entity class and the sentiment they correspond to. Some properties have different sentiments when used in context with different concepts.

Fig.1: Axioms involving the Cold class

▶ To account for this, we create axioms where these concepts are connected to certain subclasses of Entity and then assigned a positive or negative polarity. One example is the Property Cold as shown in the figure above. Cold is one of many properties for which the meaning depends on the context. The axioms within the ontology help to clarify the sentiment meaning of concepts such as Big, Dry and Funny amongst others.

▶ Sentiment is the superclass of Positive and Negative. In turn, Positive and Negative are superclasses of classes such as SustenancePositiveProperty and SustenanceNegativeProperty, respectively, which are meant to link properties with entities and their corresponding sentiment.
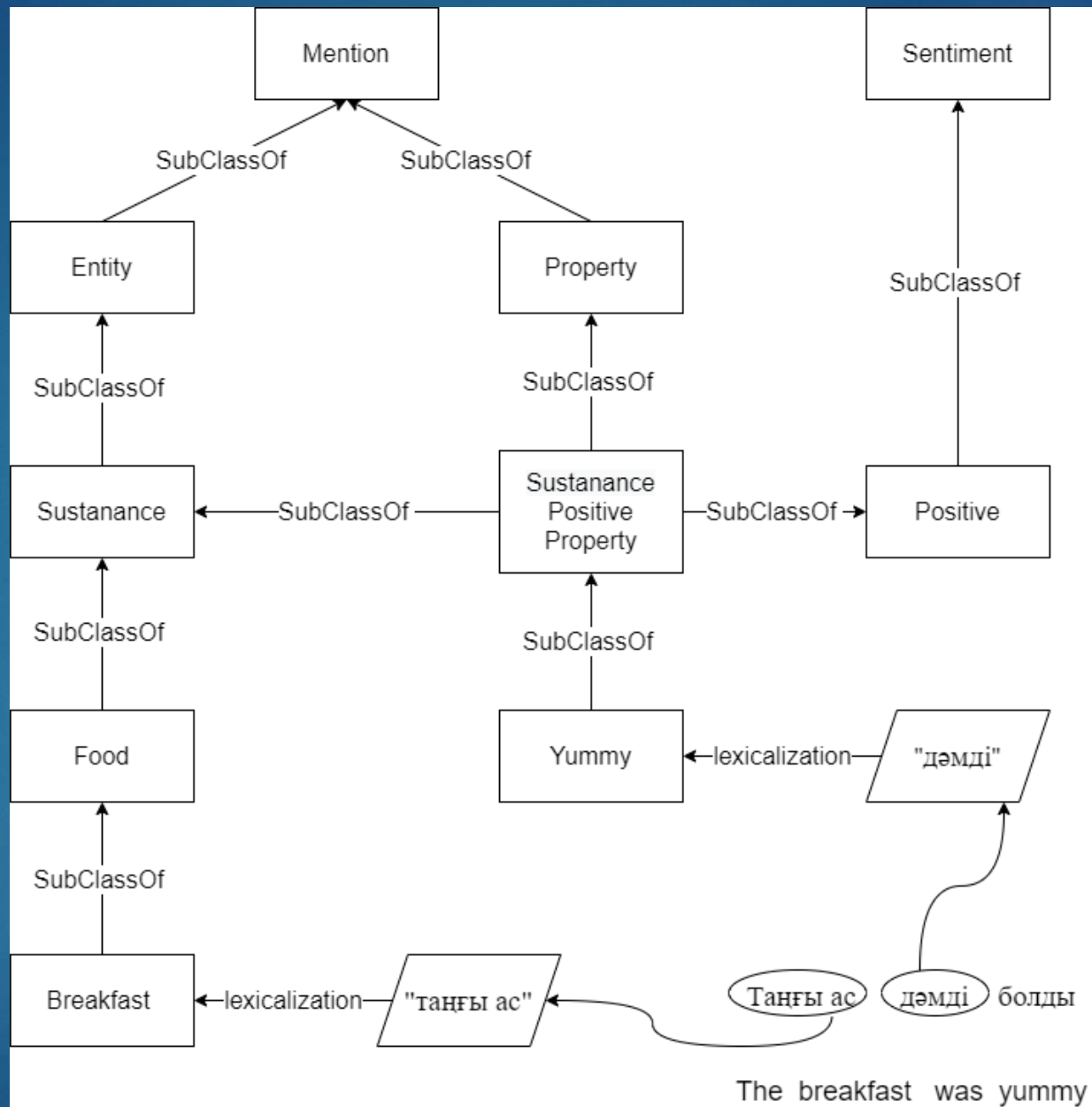
Fig. 2: Excerpt of the used ontology for the phrase "The breakfast was yummy"

- "Таңғы ас дәмді болды, алайда бөлме өте тар болды."

This is illustrated in Example above, where the aspect 'food' should have a positive sentiment where as 'ambiance' should be negative. The ontology aids the SVM in this case, because "yummy" refers to a subclass of SustenancePositiveProperty and "cramped" refers to a subclass of AmbienceNegativeProperty. The analysis for the phrase

"The breakfast was yummy" can been seen in Fig.2 .

- Note that each class is associated with multiple lexicalizations to account for different versions of spelling or for synonyms. Furthermore, the ontology is constructed manually to fit specifically with the domain of hotel reviews, using information from the training data

Conclusion
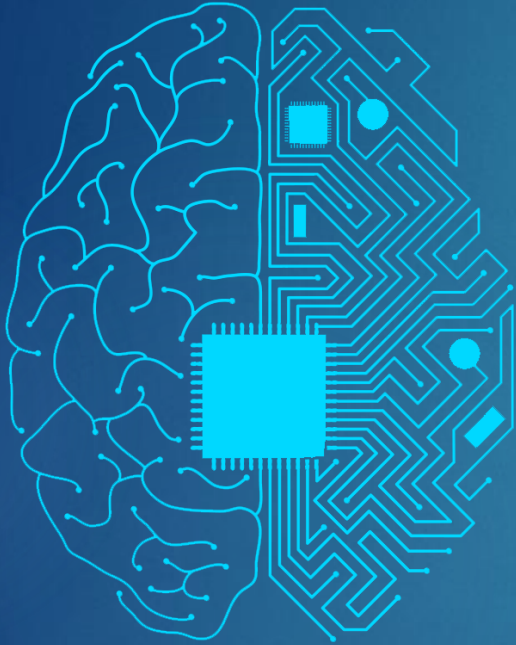
In this paper we presented an ontology approach for aspect based sentiment analysis. The ontology is constructed specifically for the domain in question. It improves the performance of the SVM for classification of aspect sentiments and reduces the reliance on training data.

Overall, the results lead us to conclude that the ontology is useful for aspect-based sentiment classification