Make up an Appetite: Recommending Personalized Menu

Michele Trevisiol
trevi@yahoo-inc.com

Luca Chiarandini
chiarluc@yahoo-inc.com

Web Research Group
Universitat Pompeu Fabra
Tànger 122-140, 08018
Barcelona, Spain

ABSTRACT

here

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

Keywords

Recommender System, Menu Builder, User Emotions

1. INTRODUCTION

In the latest years we saw a growing spread of applications which the aim is to retrieve and show users' reviews of products or services. It has been shown⌈☞ add ref.⌋ how the wisdom of the crowd allows to gain insight about the *general opinion* of a specific product/service. In this paper we are going to work with catering recommendation like restaurants, hotels,

application based on collaborative filtering reviews

have been broadly spread up becoming ones of the main sources of information⌈☞ add ref.⌋. *Yelp* and *Tripadvisor* are the top of the

Our contributions.

- We present a novel approach in restaurant recommendation, moving it to an harder and newer challenge.
- How to recognize the menu items and to build a reliable menu

  ☞ **Note**: Actually we didn't have a reliable menu.. :)

- How to infer *rates* and *emotions* on the menu items
- User-based recommender based on *menu* and *emotions*

☞ **Note**: Perché' non aggiungere related work in una sezione separata?

2. OVERVIEW

- Say how we want to proceed
- Highlight related work
- Short description of the dataset

3. ANALYSIS

In this Section we describe how we built our dataset and present some interesting results aimed at validating the soundness of the process. First of all, we will describe the dimensions along which we processed the Yelp reviews: we are interested in the sentiment, the social environment and the food. Secondly, we will analyze each dimension separately and, finally, we will combine all dimensions, highlighting interesting relationships between them.

3.1 Dimensions of the Analysis

Venue reviews could be analyzed under different points of view, since many interesting aspects could be extracted from the text of the reviews.

Since we are interested in recommending personalized menus for restaurants, *food* is certainly a dimension we would like to explore. We aim at detecting popular foods, understanding if and how food varies in terms of category of restaurant (*e.g.* Mexican, Japanese, Italian, *etc.*), and extracting frequent combination of foods, thus looking for menus.

Food alone is not enough. Reviews have a *sentiment*, that is, they could be positive reviews or negative reviews. Sentiment is therefore an important dimension to look into, especially related to the rating of the reviews (*i.e.*, stars that the person assigned to the place).

Finally, even if not essential as the previous dimensions, *social environment*, *i.e.* the people involved in the review, is an interesting dimension for the analysis. Indeed, the social context influences the behavior of people during a meal. For example, people may order different menus or have different expectations when dining with the family, with friends or with colleagues. Therefore, we expect that the content of the review may depend on this factor.

To summarize, the dimensions we will consider are:

- *Sentiment* S: amount in which the text is conceiving a positive or negative judgement;
- *Food* F: set of dishes or food present in the review;

	$S=-1$	$-1<S<+1$	$S=+1$
per review	1.22%	57.12%	41.66%
per sentence	9.87%	11.36%	78.77%

Table 1: Comparison of sentiment S in review- and sentence-level aggregation.

- *Social Environment E*: relationships between people, measured by the degree of closeness (*e.g.* a partner is closer, whereas a colleague is distant).

We will now analyze each dimension separately by highlighting how we extracted it from the text and showing interesting findings for each one of them.

3.2 Preprocessing

Since reviews are written in natural language, we perform a preprocessing aimed at reducing noise and sparsity. First of all, we split a review in sentences. For each sentence, we remove stopwords such as prepositions (*e.g.*, "to", "for", *etc.*), conjunctions (*e.g.*, "and", "or", *etc.*), pronouns and other common words. We then lemmatize all words in the sentence in order to remove the number of words and therefore the sparsity.

✍ **To-do**: Michele: check

3.3 Sentiment S

Sentiment analysis is the use of Natural Language Processing to identify subjective information from text. In this work, we are interested in understanding the *polarity* of a text, *i.e.* the amount in which it is positive or negative.

There are many method to do sentiment analysis. We will adopt a simple mechanism that relies in recognizing *polar words*, *i.e.* words that convey a positive or negative emotion, and use them to score the text. We use LIWC 2007 [5] dictionary of sentimentally-annotated words. Each word has a number of facets connected to it (*e.g.*, grammatical features, topics, *etc.*). Among all facets, we focus on those about polarity.

Given a text, we are able to score the positiveness by counting the occurrence of polarity words in it. The sentiment score of a text is given by Equation 1:

$$S = \frac{p-n}{p+n} \tag{1}$$

where p is the number of positive words, and n is the number of negative words.

3.3.1 Sentence- vs. Review-level Sentiment Analysis

The first question we ask is whether dividing the text of the reviews in sentences affects the results. To answer this, we compare the distribution of sentiment in the whole review and the one of its sentences.

Table 1 shows the comparison between the two cases. We show three cases: 1. *negative*, when $S=-1$; 2. *mixed*, when $0<S<1$; and 3. *positive*, when $S=1$. We can see that, in the case of sentiment of sentences, there is a majority of positive and negative, while review sentiments are more mixed. Indeed, mixed sentiments occur only in 12% of the sentences, against 43% in the case of review-level aggregation.

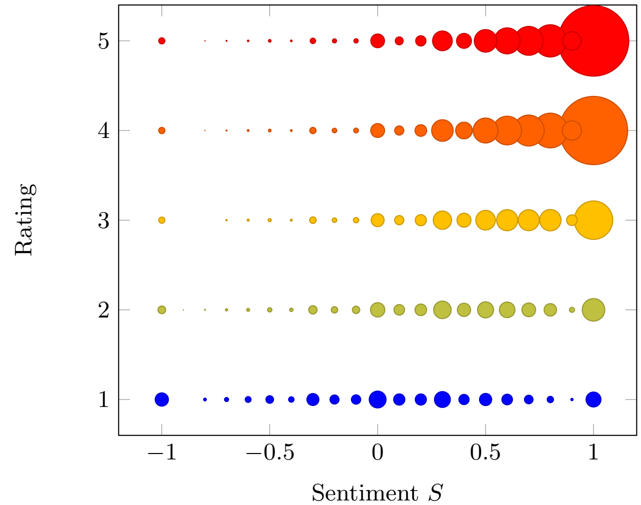

Figure 1: Comparison between sentiment S and rating R (expressed in number of stars, 1 is worse, 5 is best). The area of the circle represents the amount of reviews.

We conclude that splitting by sentence allows us to get a more precise, clean, and localized characterization of the text.

3.3.2 Sentiment and Ratings

It is natural to expect that the sentiment of a review is related to the ratings given by users. Figure 1 shows the distribution of sentiments for each rating. We can see that the amount of positive reviews decreases and that the amount of negative reviews decreases with the rating. The tendency is towards positive reviews, which is a very well known effect called *Pollyanna principle* [2].

A small amount of purely negative reviews ($S=-1$) are always present, even for highly-rated reviews. Observing such reviews manually, we understand that they are not related to the place itself, but rather to contingencies. For example, the following reviewer is disappointed that a cafe is closed, but gave the maximum review anyway:

> [...] I stopped by two days ago unaware that they had closed. I am severely bummed. This place is irreplaceable! [...]

3.4 Food F

Food, along with the quality of service, is the most important aspect of restaurant and cafe reviews. People write about food in their reviews and often share their favorite menu at a particular place. Extracting food from text is a difficult task due to the large amount of ingredients and local expressions.

We implemented a basic method which captures the most popular foods in reviews based on a dictionary extracted from three publicly available sources:

- *Oregon State University Food Glossary*[1]: this is a multi-language glossary of food which contains ingredients as well as scientific names. Using a web-crawler, we built

[1] `http://food.oregonstate.edu/`

Food	
pizza	3.38%
salad	2.49%
chicken	2.43%
sandwich	2.39%
burger	2.23%
beer	1.89%

Table 2: Most frequent foods in the dataset.

American	
salad	2.89%
beer	2.55%
wine	2.35%
burger	2.21%
sandwich	1.82%
chicken	1.79%
cheese	1.77%
dessert	1.73%
bread	1.46%
sauce	1.02%
meat	0.93%

Italian	
pizza	6.93%
salad	3.63%
wine	3.37%
pasta	2.77%
bread	2.70%
sauce	2.50%
cheese	2.21%
dessert	1.94%
sandwich	1.87%
chicken	1.66%
bruschetta	1.21%

Mexican	
taco	6.00%
salsa	5.88%
chip	4.17%
burrito	3.44%
bean	2.72%
margarita	2.72%
chicken	2.63%
cheese	2.28%
rice	2.15%
enchilada	2.09%
tortilla	1.81%

Chinese	
chicken	6.11%
beef	3.08%
shrimp	2.63%
soup	2.47%
sauce	2.40%
fried rice	2.29%
noodle	2.00%
spicy	1.92%
pork	1.90%
egg roll	1.70%
rice	1.67%

Table 3: Most frequent foods for various type of restaurants.

a dictionary based on the titles of the food pages in the glossary.

- *WordNet*[2]: WordNet is a large lexical database of English. We built a dictionary containing all nouns in the "food" group.

- *BBC Food*[3]: BBC Food is a web portal of recipes and ingredients. It contains a large amount of recipes written in English. We crawled the pages and extracted all ingredients and recipes. For each page, we crawled the displayed image and the description. Also, for each recipe, we were able to crawl all its ingredients. The final dictionary consists of around 9000 items.

After building the dictionary, we lemmatize all words, we manually remove some noise and we find such words in the text of the reviews.

3.4.1 Statistics

Table 2 shows the most frequent food words in the dataset, alongside with the percentage of occurrences. A breakdown by type of restaurant is provided in Table 3. We can see

[2] `http://wordnet.princeton.edu/`
[3] `http://www.bbc.co.uk/food/`

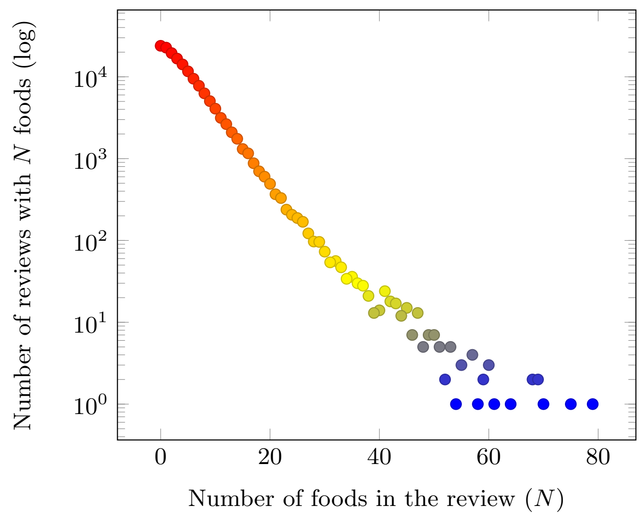

Figure 2: Number of food per reviews.

that the food which appear in the list are indeed typical of the particular cuisine.

As for the coverage of the foods in the reviews, on average, we detect 4.7 food words per review. Figure 2 shows the head of the distribution of food per review. The number of food per review is uniform across ratings and types of restaurants.

3.5 Social Environment E

✍ **To-do**: How to extract the dimensions from reviews

✍ **To-do**: Show validity of feature by plotting distributions. Show examples

3.6 Food and Sentiment

Having inspected every dimension by itself, it is now time to analyze dimensions jointly. We start by analyzing maybe the most interesting one for our goal: sentiment and food.

3.6.1 Sentence-level Food Sentiment

We observed in Section 3.3.1 that extracting sentiment from sentences gives a more localized and clear signal than extracting it from the whole review text. In addition, being able to detect the sentiment of particular sentences in the reviews allows us to better connect the sentiment to the food words. It is indeed quite common (57% of reviews, see Table 1) for people to write a mixed-sentiment reviews. This often happens when reviewing more than one dish, as for example:

> [...] Pizza crust & toppings are excellent. However the pizza sauce was too salty. [...]

Bearing this in mind, we assign to each occurrence of a food word the sentiment S of the sentence it belongs.

In order to evaluate the way in which we assign sentiment to food words, we compare it to the ratings given by users. We build the ranking of food words based solely on the sentence-level sentiments r_S, and we compare it to the ranking we would obtain using the ratings of the reviews r_R.