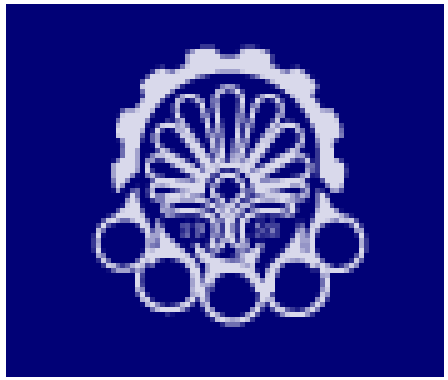
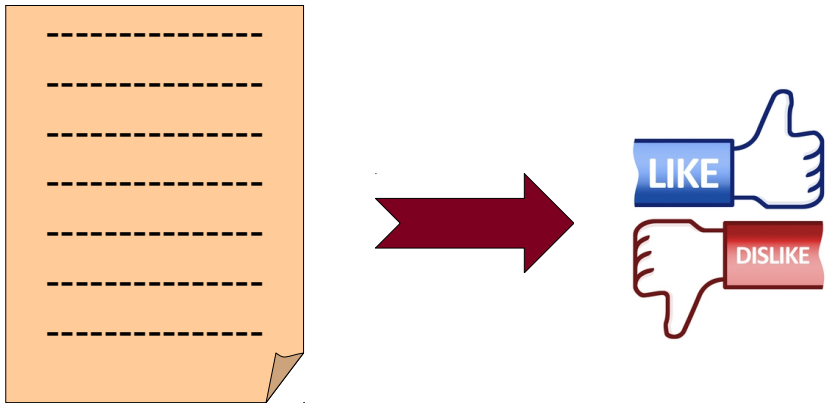


Statistical Natural Language Processing

Lecture 13: Sentiment Analysis

Dr. Momtazi
Amirkabir University of Technology

Sentiment Analysis



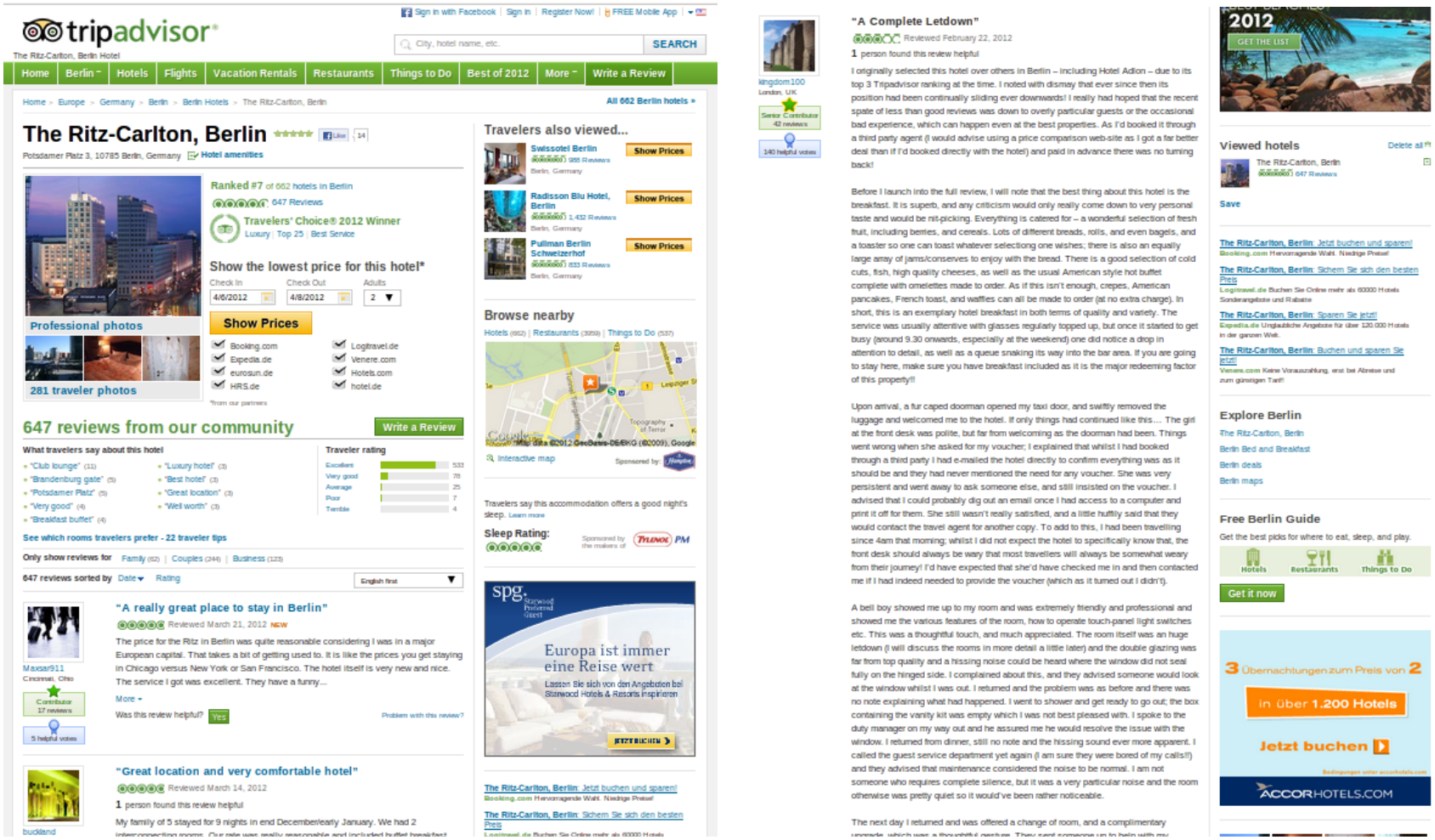
Outline

1. Applications
2. Task
3. Machine Learning Approach
4. Rule-based Approach

Outline

1. Applications
2. Task
3. Machine Learning Approach
4. Rule-based Approach

Hotel Reviews



"A Complete Letdown"

Reviewed February 22, 2012

1 person found this review helpful

I originally selected this hotel over others in Berlin – including Hotel Adlon – due to its top 3 Tripadvisor ranking at the time. I noted with dismay that ever since then its position had been continually sliding ever downwards! I really had hoped that the recent spate of less than good reviews was down to overly particular guests or the occasional bad experience, which can happen even at the best properties. As I'd booked it through a third party agent (I would advise using a price comparison web-site as I got a far better deal than if I'd booked directly with the hotel) and paid in advance there was no turning back!

Before I launch into the full review, I will note that the best thing about this hotel is the breakfast. It is superb, and any criticism would only really come down to very personal taste and would be nit-picking. Everything is catered for – a wonderful selection of fresh fruit, including berries, and cereals. Lots of different breads, rolls, and even bagels, and a toaster so one can toast whatever selectiong one wishes; there is also an equally large array of jams/conserves to enjoy with the bread. There is a good selection of cold cuts, fish, high quality cheeses, as well as the usual American style hot buffet complete with omelettes made to order. As if this isn't enough, crepes, American pancakes, French toast, and waffles can all be made to order (at no extra charge). In short, this is an exemplary hotel breakfast in both terms of quality and variety. The service was usually attentive with glasses regularly topped up, but once it started to get busy (around 9.30 onwards, especially at the weekend) one did notice a drop in attention to detail, as well as a queue snaking its way into the bar area. If you are going to stay here, make sure you have breakfast included as it is the major redeeming factor of this property!!

Upon arrival, a fur caped doorman opened my taxi door, and swiftly removed the luggage and welcomed me to the hotel. If only things had continued like this… The girl at the front desk was polite, but far from welcoming as the doorman had been. Things went wrong when she asked for my voucher; I explained that whilst I had booked through a third party I had e-mailed the hotel directly to confirm everything was as it should be and they had never mentioned the need for any voucher. She was very persistent and went away to ask someone else, and still insisted on the voucher. I advised that I could probably dig out an email once I had access to a computer and print it off for them. She still wasn't really satisfied, and a little huffily said that they would contact the travel agent for another copy. To add to this, I had been travelling since 4am that morning; whilst I did not expect the hotel to specifically know that, the front desk should always be wary that most travellers will always be somewhat weary from their journey! I'd have expected that she'd have checked me in and then contacted me if I had indeed needed to provide the voucher (which as it turned out I didn't).

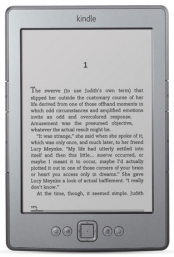
A bell boy showed me up to my room and was extremely friendly and professional and showed me the various features of the room, how to operate touch-panel light switches etc. This was a thoughtful touch, and much appreciated. The room itself was an huge letdown (I will discuss the rooms in more detail a little later) and the double glazing was far from top quality and a hissing noise could be heard where the window did not seal fully on the hinged side. I complained about this, and they advised someone would look at the window whilst I was out. I returned and the problem was as before and there was no note explaining what had happened. I went to shower and get ready to go out; the box containing the vanity kit was empty which I was not best pleased with. I spoke to the duty manager on my way out and he assured me he would resolve the issue with the window. I returned from dinner, still no note and the hissing sound ever more apparent. I called the guest service department yet again (I am sure they were bored of my calls!!) and they advised that maintenance considered the noise to be normal. I am not someone who requires complete silence, but it was a very particular noise and the room otherwise was pretty quiet so it would've been rather noticeable.

The next day I returned and was offered a change of room, and a complimentary

Value
Location
Sleep Quality

Rooms
Cleanliness
Service

Product Reviews



(4,810)
4.3 out of 5 stars

5 star		3,120
4 star		884
3 star		293
2 star		159
1 star		354

Picture Quality
Ease of Use
Size
Weight
Color
Zoom

Social Media

Event Analysis and Prediction

- Analyzing the side effects of events in different communities
- Predicting the election results
- Predicting the Stock exchange
- ...

Outline

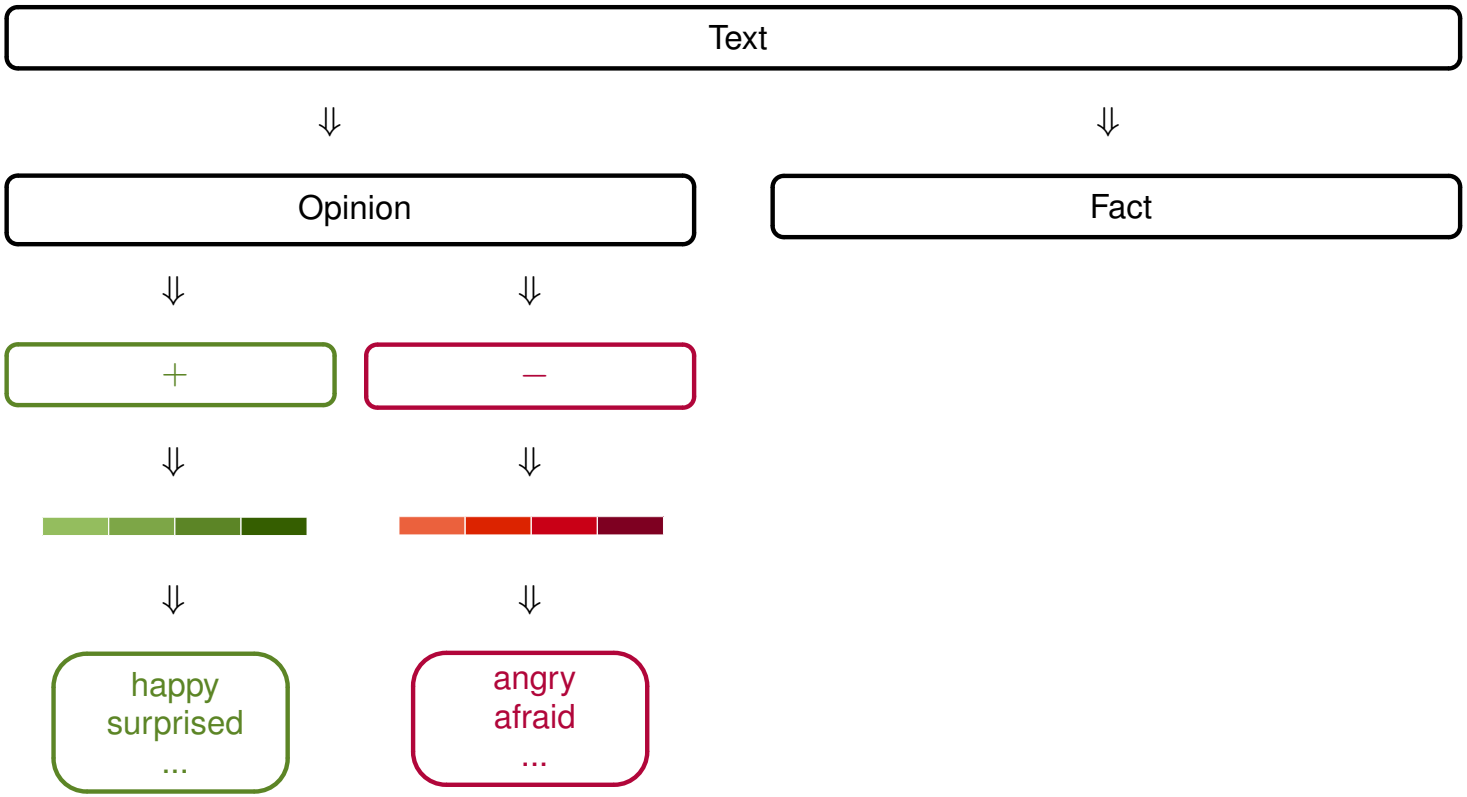
1 Applications

2 Task

3 Machine Learning Approach

4 Rule-based Approach

Sentiment Analysis Levels



Sentiment Analysis Tasks

- Opinion holder detection
- Opinion target detection
- Aspect detection

Students like Wikipedia because it is easy to use and it sounds authoritative.
op holder *target*

I had a nice stay in this hotel and the rooms were very clean.
. *aspect*

The restaurant has an amazing view but the staffs are unfriendly.
. *aspect* *aspect*

Mixed opinions

Other Names

- Opinion mining
- Opinion extraction
- Sentiment mining
- Subjectivity detection
- Subjectivity analysis

Sentiment Analysis Approaches

- Machine learning methods
  $\Rightarrow$classification

- Rule-based methods
  $\Rightarrow$ dictionary oriented

Outline

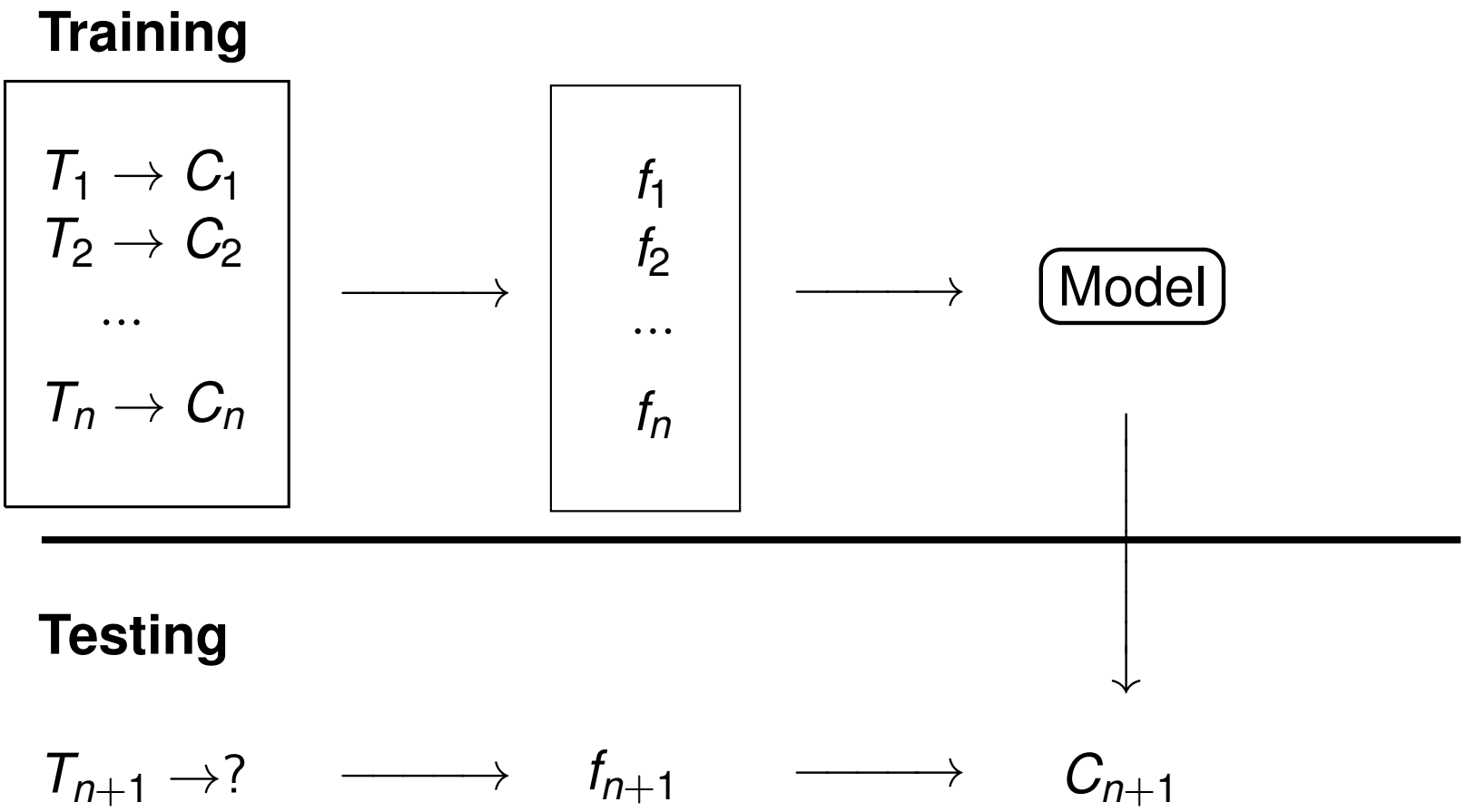
1 Applications

2 Task

3 Machine Learning Approach

4 Rule-based Approach

Machine Learning Approach



Sentiment Classification

- Using any kinds of supervised classifiers
  - *K* Nearest Neighbor
  - Support Vector Machines
  - Naïve Bayes
  - Maximum Entropy
  - Logistic Regression
  - ...

Features

- Word
  - All words or adjectives?
    All words works better than adjectives only
  - Word occurrence or frequency?
    Word occurrence is more useful than frequency
    - Using binary value for words
    - Replace all word counts higher than 0 in each text by 1

Features

- Negation
  - Negation words change the text polarity
    - Adding prefix *NOT_* to every word between negation and next punctuation

"I did not like the restaurant location, but the food ..."

I did not NOT-like NOT-the NOT-restaurant NOT-location but the food ...

Features

- Other emotions
  - Considering emoticons as additional features
    - :)
    - :(

Fine-grained Analysis

- Dealing with finer classes of sentiment
  -3,-2,-1,+1,+2,+3

- Approaches
  - Using multiclass classifier (6 classes in this case)
  - Using two level classifier
    - First level: polarity classifier (positive or negative)
    - Second level: strength classifier (1 or 2 or 3)

Outline

1 Applications

2 Task

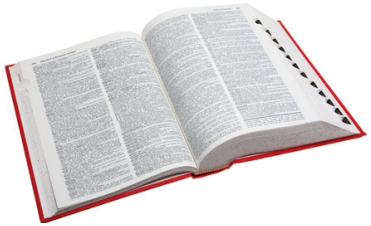
3 Machine Learning Approach

4 Rule-based Approach

Rule-based Approach

Training

$T_1 \rightarrow C_1$
$T_2 \rightarrow C_2$
...
$T_n \rightarrow C_n$



good	bad
love	hate
brave	lie
intelligent	ugly
nice	poor
...	...

Testing

$T_{n+1} \rightarrow ?$ $\longrightarrow$ C_{n+1}

Rule-based Approach

- Looking for opinionated words in each text
- Classifying the text based on the number of positive and negative words

- Considering different rules for classification
  - Fine-grained dictionary
  - Negation words
  - Booster words
  - Idioms
  - Emoticons
  - Mixed opinions
  - Linguistic features of the language

Rule-based Approach

- Fine-grained Dictionary

"It was a good song."

"The song was excellent."

Rule-based Approach

- Negation Words

“The song was good.”

“The song was not good.”

Rule-based Approach

- Booster Words

"The song was interesting."

"The song was very interesting."

"The song was somewhat interesting."

Rule-based Approach

- Idioms

“shock horror”

Rule-based Approach

- Mixed Opinions

"The song was good, but I think its title was strange."

Rule-based Approach

- Linguistic Features
  - The position of negation or boosting words differ from one language to another

"I do not love the song."

Opinion Dictionary

- English
  - Subjectivity Clues (2005)
  - SentiSpin (2005)
  - SentiWordNet (2006)
  - Polarity Enhancement (2009)
  - SentiStrength (2010)

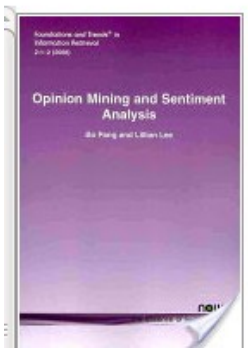
Machine Learning with Opinion Dictionary

- Using opinion words as a feature in the algorithms
- Ignoring other words in the text

Adjectives alone do not work well,
but opinion words are the best features to be used

Further Reading

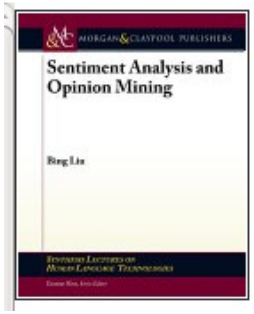
Opinion Mining and Sentiment Analysis



Bo Pang, Lillian Lee

Now Publishers Inc, 2008 - Computers - 137 pages

Sentiment Analysis and Opinion Mining



Bing Liu

Morgan & Claypool Publishers, Ordibehesht 12, 1391 AP - Computers - 167 pages