
Sentiment Analysis and other non-explicit semantics of text

(using some material from Dan Jurafsky)

Why sentiment analysis?

- *Movie*: is this review positive or negative?
- *Products*: what do people think about the new iPhone?
- *Customer relations*: what do people say about your company?
- *Public sentiment*: how is consumer confidence? Is despair increasing?
- *Politics*: what do people think about this candidate or issue?
- *Prediction*: predict election outcomes or market trends from sentiment

Types of tasks:

- Positive or negative movie review?



- unbelievably disappointing



- Full of zany characters and richly applied satire, and some great plot twists



- this is the greatest screwball comedy ever filmed



- It was pathetic. The worst part about it was the boxing scenes.

Types of Tasks: Google Product Search



HP Officejet 6500A Plus e-All-in-One Color Ink-jet - Fax / copier / printer / scanner

\$89 online, \$100 nearby ★★★★★ 377 reviews

September 2010 - Printer - HP - Inkjet - Office - Copier - Color - Scanner - Fax - 250 shi

Reviews

Summary - Based on 377 reviews



What people are saying

ease of use	<div><div></div><div></div></div>	"This was very easy to setup to four computers."
value	<div><div></div><div></div></div>	"Appreciate good quality at a fair price."
setup	<div><div></div><div></div></div>	"Overall pretty easy setup."
customer service	<div><div></div><div></div></div>	"I DO like honest tech support people."
size	<div><div></div><div></div></div>	"Pretty Paper weight."
mode	<div><div></div><div></div></div>	"Photos were fair on the high quality mode."
colors	<div><div></div><div></div></div>	"Full color prints came out with great quality."

Types of tasks: Bing Shopping

HP Officejet 6500A E710N Multifunction Printer

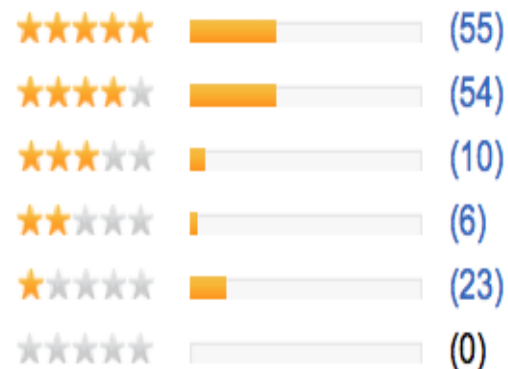
[Product summary](#) [Find best price](#) **Customer reviews** [Specifications](#) [Related items](#)



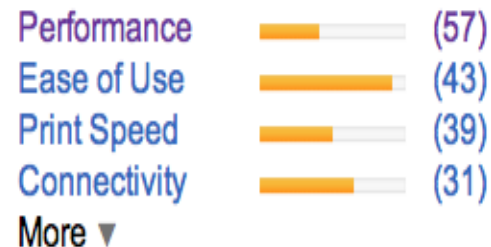
\$121.53 - \$242.39 (14 stores)

☐ Compare

Average rating ★★★★★ (144)



Most mentioned



Show reviews by source

[Best Buy \(140\)](#)
[CNET \(5\)](#)
[Amazon.com \(3\)](#)

Target Sentiment on Twitter

- [Twitter Sentiment App](#)
- Alec Go, Richa Bhayani, Lei Huang.
2009. Twitter Sentiment Classification
using Distant Supervision

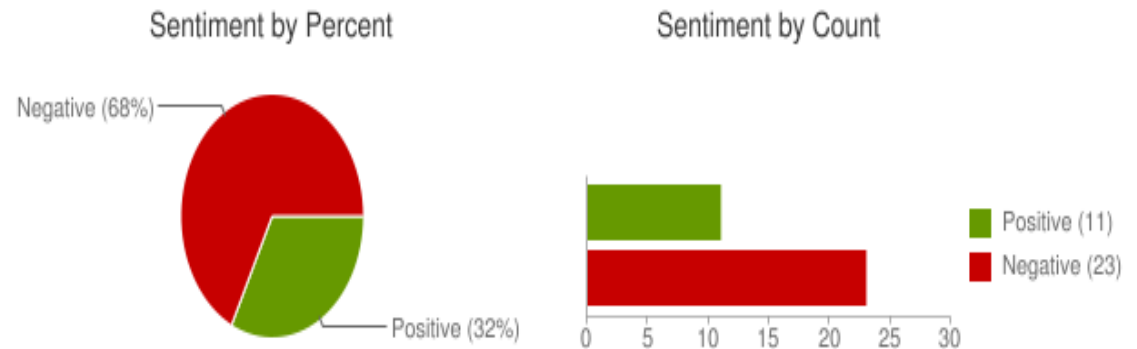
Type in a word and we'll highlight the good and the bad

"united airlines"

Search

[Save this search](#)

Sentiment analysis for "united airlines"



[jacobson](#): OMG... Could **@United airlines** have worse customer service? W8g now 15 minutes on hold 4 questions about a flight 2DAY that need a human.
[Posted 2 hours ago](#)

[12345clumsy6789](#): I hate **United Airlines** Ceiling!!! Fukn impossible to get my conduit in this damn mess! ?
[Posted 2 hours ago](#)

[EMLandPRGbelgiu](#): EML/PRG fly with Q8 **united airlines** and 24seven to an exotic destination. <http://t.co/Z9QloAjF>
[Posted 2 hours ago](#)

[CountAdam](#): FANTASTIC customer service from **United Airlines** at XNA today. Is tweet more, but cell phones off now!
[Posted 4 hours ago](#)

Terminology

- For the more subtle aspects of the semantics of text:
- **Affective aspects of text** is that which is “influenced by or resulting from emotions”
 - One aspect of non-factual aspects of text
- **Subjective aspects of text**
“The **linguistic** expression of somebody’ s **opinions**, **sentiments**, emotions, evaluations, beliefs, speculations (*private states*)”
 - A private state is not open to objective observation or verification
 - Subjectivity analysis would classify parts of text as to whether it was subjective or objective

Scherer Typology of Affective States

- **Emotion:** brief organically synchronized ... evaluation of a major event
 - *angry, sad, joyful, fearful, ashamed, proud, elated*
- **Mood:** diffuse non-caused low-intensity long-duration change in subjective feeling
 - *cheerful, gloomy, irritable, listless, depressed, buoyant*
- **Interpersonal stances:** affective stance toward another person in a specific interaction
 - *friendly, flirtatious, distant, cold, warm, supportive, contemptuous*
- **Attitudes:** enduring, affectively colored beliefs, dispositions towards objects or persons
 - *liking, loving, hating, valuing, desiring*
- **Personality traits:** stable personality dispositions and typical behavior tendencies
 - *nervous, anxious, reckless, morose, hostile, jealous*

Sentiment Analysis

- **Emotion:** brief organically synchronized ... evaluation of a major event
 - *angry, sad, joyful, fearful, ashamed, proud, elated*
- **Mood:** diffuse non-caused low-intensity long-duration change in subjective feeling
 - *cheerful, gloomy, irritable, listless, depressed, buoyant*
- **Interpersonal stances:** affective stance toward another person in a specific interaction
 - *friendly, flirtatious, distant, cold, warm, supportive, contemptuous*
- **Attitudes: enduring, affectively colored beliefs, dispositions towards objects or persons**
 - *liking, loving, hating, valuing, desiring*
- **Personality traits:** stable personality dispositions and typical behavior tendencies
 - *nervous, anxious, reckless, morose, hostile, jealous*

Sentiment Analysis

- Sentiment analysis is the detection of attitudes
 - “enduring, affectively colored beliefs, dispositions towards objects or persons”
 - **Holder (source)** of attitude
 - **Target (aspect)** of attitude
 - **Type** of attitude
 - From a set of types
 - Like, love, hate, value, desire, etc.
 - Or (more commonly) simple weighted polarity:
 - Positive, negative, neutral, together with strength
 - **Text** containing the attitude
 - Sentence or entire document

Sentiment Analysis Task Levels

- **Simplest task is polarity:**
 - Is the attitude of this text positive or negative?
 - Negative / positive attitude of reporter / blogger
 - Favorable / unfavorable review of a product
 - Right / left political leaning of speaker
- **More complex:**
 - Rank the attitude of this text from 1 to 5
 - Sometimes called strength or intensity
- **Advanced:**
 - Detect the target, source or complex attitude type
 - May also be referred to as
opinion extraction, opinion mining, or sentiment mining

What's the problem?

- Consider classifying a subjective text unit as either positive or negative.
 - Example: The most thoroughly joyless and inept film of the year, and one of the worst of the decade. [Mick LaSalle, describing *Gigli*]
- Can't we just look for words like *great* or *terrible* ?
 - Yes, but ...
 - This laptop is a **great deal**.
 - A **great deal** of media attention surrounded the release of the new laptop.
 - This laptop is a **great deal** ... and I've got a nice bridge you might be interested in.
 - Words have different meanings in different contexts

What's the problem?

- Subtlety, sarcasm or metaphor:
 - Perfume review in *Perfumes: the Guide*:
 - “If you are reading this because it is your darling fragrance, please wear it at home exclusively, and tape the windows shut.”
 - Dorothy Parker on Katherine Hepburn
 - “She runs the gamut of emotions from A to B”

Thwarted Expectations and Ordering Effects

- “This film should be brilliant. It sounds like a great plot, the actors are first grade, and the supporting cast is good as well, and Stallone is attempting to deliver a good performance. However, it **can’t hold up.**”
- Well as usual Keanu Reeves is nothing special, but surprisingly, the very talented Laurence Fishbourne is **not so good** either, I was surprised.

Domain Adaptation

- Certain sentiment-related indicators seem domain-dependent.
 - .Read the book.: good for book reviews, bad for movie reviews
 - .Unpredictable.: good for movie plots, bad for a car's steering [Turney '02]
- In general, sentiment classifiers (especially those created via supervised learning) have been shown to often be domain dependent
 - [Turney '02, Engström '04, Read 05, Aue & Gamon '05, Blitzer, Dredze & Pereira '07].