

Natural Language Processing

Sentiment Analysis

Agenda

- Why Sentiment Analysis
- What is Sentiment Analysis
- Sentiment Analysis – Application Areas
- Sentiment Analysis – How Far
- Sentiment analysis – Using TextBlob
- Working of Classification
- Sentiment analysis – Levels

“What people think”

What others think has always been an important piece of information

“Which car should I buy?”

“Which schools should I apply to?”

“Which Professor to work for?”

“Whom should I vote for?”



“So, whom shall I ask?”

- Pre Web

- Friends and relatives
- Acquaintances
- Consumer Reports



- Post Web

“...I don't know who..but apparently it's a good phone. It has good battery life and...”

- Blogs (google blogs, livejournal)
- E-commerce sites (amazon, ebay)
- Review sites (CNET, PC Magazine)
- Discussion forums (forums.craigslist.org, forums.macrumors.com)
- Friends and Relatives (occasionally)



“Whoala! I have the reviews I need”

*Now that I have “**too much**” information on one topic...I could easily form my opinion and make decisions...*

Is this true?

...Not Quite

Searching for reviews may be difficult

Can you search for opinions as conveniently as general Web search?

eg: is it easy to search for “*iPhone vs Google Phone*”?

Motivation

- Knowing sentiment is a very natural ability of a human being.
 - Can a machine be trained to do it?
- Sentiment Analysis aims at getting sentiment-related knowledge especially from the huge amount of information on the internet
- Can be generally used to understand opinion in a set of documents

Sentiment Analysis

- Sentiment Analysis is the process of computationally determining whether a piece of content is positive, negative or neutral.
- It is also known as Opinion Mining.
- Indicates
 - The emotional effect the text or document has on the reader (or the author wants to convey).
 - the attitude of the author about the subject.

Sentiment Analysis

- Positive words/sentences have a positive sentiment attached to them
 - text indicates happiness, enthusiasm, kindness etc.,
 - generally classified as having a positive sentiment
- Negative sentences have a negative sentiment attached to them
 - text indicates sadness, hate, violence, discrimination etc.,
 - generally classified as having a negative sentiment
- When no emotions are implied, they're classified as neutral

Sentiment Analysis

- is a Natural Language Processing and Information extraction task
- aims to obtain writer's feelings expressed in positive or negative comments, questions and requests, by analyzing a large numbers of documents.
- Generally speaking, sentiment analysis aims to determine the attitude of a speaker or a writer with respect to some topic or the overall tonality of a document.

Why Sentiment Analysis ?

- Sentiment Analysis is mainly used to gauge the views of public regarding any action, event, person, policy or product.
- It has become a very potent weapon even for politicians to assess the public reaction over their statements.
- These days Opinion Mining has reached an advanced stage where several outcomes can be predicted using large datasets and machine learning etc.

Challenges

- Contrasts with standard text-based categorization
- Domain dependent
- Sarcasm
- Thwarted expressions

the sentences/words that contradict the overall sentiment of the set are in majority

Example: The actors are good, the music is brilliant and appealing. Yet, the movie fails to strike a chord.

Sentiment Analysis – also called

- Opinion mining
 - Sentiment analysis
 - Sentiment mining
 - Subjectivity detection
 - ...
-
- Often used synonymously
 - Some shadings in meaning
 - “sentiment analysis” describes the current mainstream task best

Where - Sentiment Analysis

- *Movie*: is this review positive or negative?
- *Products*: what do people think about the new iPhone?
- *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



ON OFFER



[^ Back to top](#)

Special price if bought with this item

1 Item ₹10,999 + 2 Add-ons ₹418 = Total ₹11,417

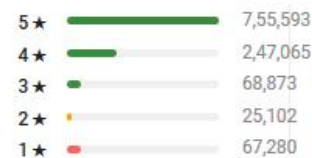
[ADD 3 ITEMS TO CART](#)

Ratings and Reviews

[Rate this product](#)

4.4 ★

11,63,913
Ratings &
2,09,533
Reviews



Camera



Battery



Display



Value for Money

5 ★ Awesome

Superb phone ..
Built quality excellent..
Jio runs superb ... very fast
Small disappointment with back camera ..
Front camera is ok..
Sleek design....
Out of 64gb 52 gb is available..
Hats off for the fast delivery by flipkart..
Overall satisfied with the phone...

Raja Sekhar Certified Buyer 28 Jan, 2017

49256

7700

5 ★ Best in the market!

My review in abbreviation!

sentiment analysis - (buzzillions.com)

4. On Stage MA100 - Screw Adapter



★★★★★ 4.5 read 22 reviews

"Perfect" – Juanpianoman

"Super Useful Adapter Hard to Find!" – Mike O.

“Good for attaching to boom pole, BU...” – Russell Rules

starting at **1,95 USD**

+ Compare

5. Verbatim Forecast Bimini 4 Piece Value Set



2.6 read 12 reviews

"I bought this luggage set ..." – RickHawaii

"bum wheels" – bumed1

"I purchase the identical ..." – whoopingcough

starting at 1,99 USD

[+ Compare](#)

6. Pro-Co Sound Rack Screws and Washers



★★★★★ 5.0 [read 2 reviews](#)

"What can you say "they are screws"" – JM

“hard to mess up a screw” – Gary

starting at **2,50 USD**

+ Compare

Aggregations - buzzillions.com

Compare Products					
The following products are listed below	Advanced amateur	—	1%	—	—
	Beginner	Needed it to use external	—	—	2%
	Serious amateur	Great access to all ports	—	—	2%
	Casual photographer	Long lasting	—	—	81%
	Student	Fast charge	—	—	34%
	Budget shopper	Good value	—	—	17%
	Quality oriented	Good power output	—	—	12%
	Pros				
	Easy to use	No worries about dead bat	—	—	2%
	Durable	Good long lasting	—	—	2%
Avg. Pros	Reliable	Super compact size	—	—	2%
	Cons				
	Great picture quality	Heavy / bulky	—	34%	7%
	Well built / quality	Pricy / poor value	—	5%	3%
	Fast / accurate auto-focus	Heavy bulky	—	3%	—
Ratings		Price	—	2%	2%
		Weight	—	2%	—
		Heavy bulky	—	1%	—
		Cost	—	1%	—
User Type					

Sentiment analysis



The image shows a promotional graphic for Sendible. On the left, a dark blue background features the Sendible logo at the top. Below it, the text reads "The #1 Social Media Management Tool for Agencies". Further down, it says "Sendible is simply the *best* way for teams to manage social media for multiple clients — and get results." At the bottom of this section are two buttons: "START YOUR FREE TRIAL" and "SOLUTIONS FOR AGENCIES". On the right, a laptop displays the Sendible web application interface. The interface includes a top navigation bar with links for FEATURES, PRICING, GET A DEMO, BLOG, LOGIN, and SIGN UP FOR FREE. The main dashboard shows a list of social media posts from various users, including Don Pizzaro, Esha Sundrani, Jay Perkins, Entrepreneur Foundry, Patrick Whatman, and James Jane. The interface also includes a left sidebar with navigation options like INBOX, OUTBOX, and TASKS, and a right sidebar with a search bar and a list of social media accounts.

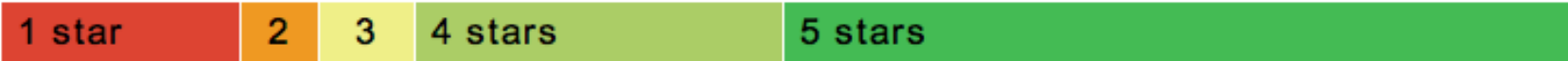
Google Product Search



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 sh

Reviews

Summary - Based on 377 reviews




What people are saying

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

Google Product Search


HP Officejet 6500A E710N Multifunction Printer


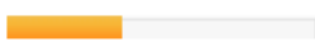

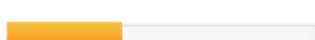

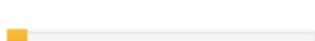






[Product summary](#) [Find best price](#) **Customer reviews** [Specifications](#) [Related items](#)







\$121.53 - \$242.39 (14 stores)

☐ Compare

Average rating  (144)

		(55)
		(54)
		(10)
		(6)
		(23)
		(0)

Most mentioned

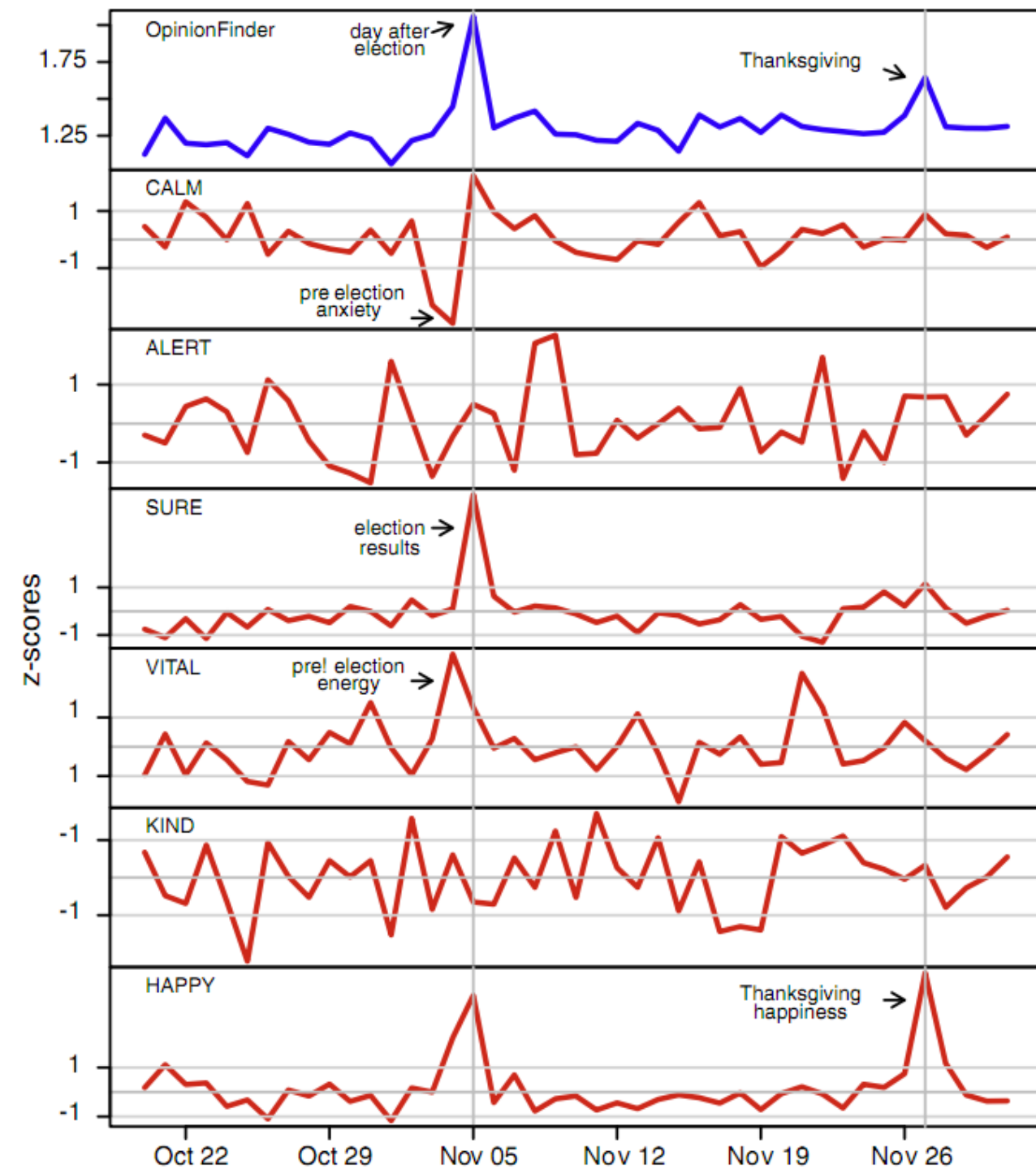
Performance		(57)
Ease of Use		(43)
Print Speed		(39)
Connectivity		(31)
More ▼		

Show reviews by source

- Best Buy (140)
- CNET (5)
- Amazon.com (3)

Twitter sentiment:

Johan Bollen, Huina Mao, Xiaojun Zeng. 2011.
[Twitter mood predicts the stock market](#),
Journal of Computational Science 2:1, 1-8.
10.1016/j.jocs.2010.12.007.



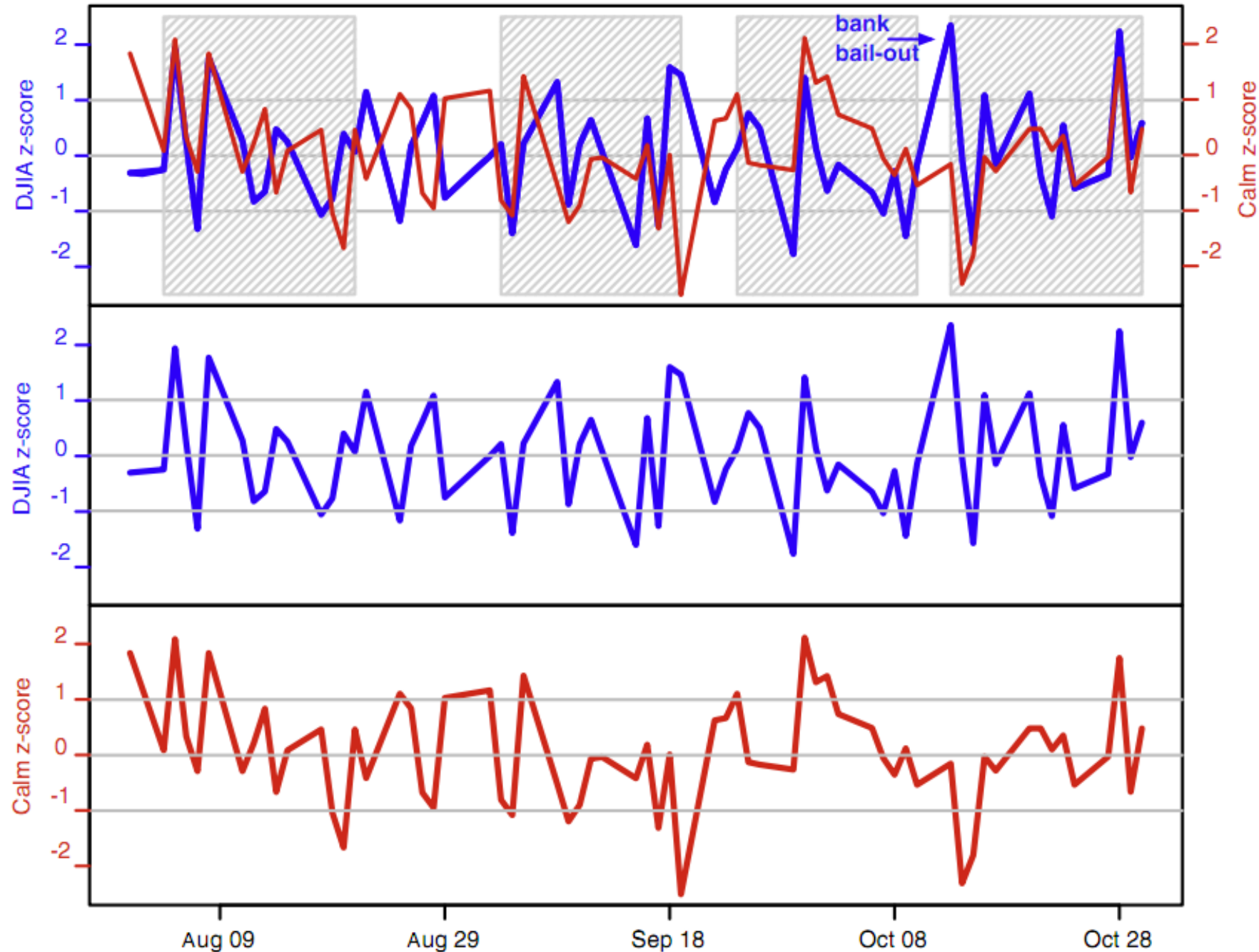
Bollen et al. (2011)

CALM predicts DJIA
3 days later

At least one current
hedge fund uses
this algorithm

Dow Jones

CALM



Target Sentiment on Twitter

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

"united airlines"

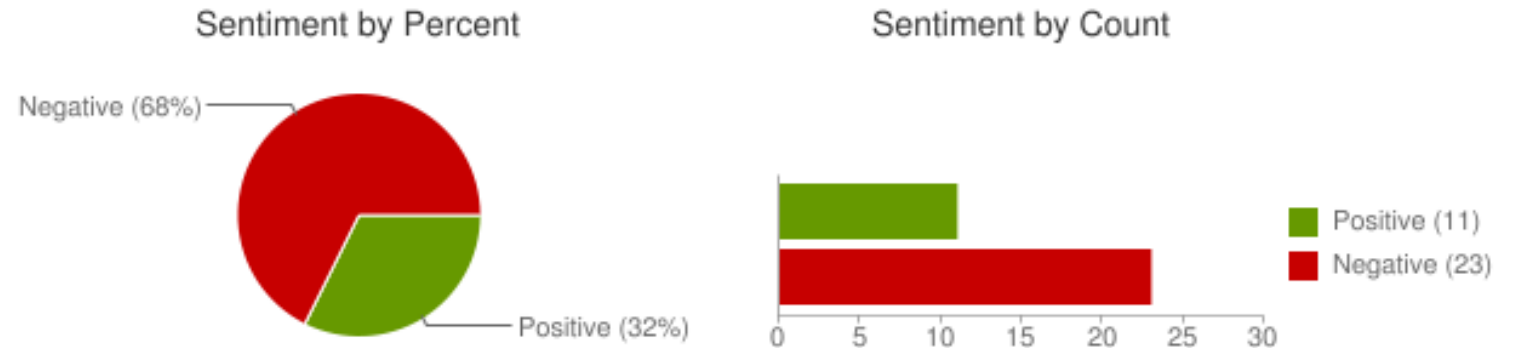
Search

[Save this search](#)

Twitter Sentiment App

Alec Go, Richa Bhayani, Lei Huang. 2009.
Twitter Sentiment Classification using
Distant Supervision

Sentiment analysis for "united airlines"



ljacobson: 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](#)

Application Areas Summarized

- **Businesses and organizations:** interested in opinions
 - product and service benchmarking
 - market intelligence
 - survey on a topic
- **Individuals:** interested in other's opinions when
 - Purchasing a product
 - Using a service
 - Tracking political topics
 - Other decision making tasks
- **Ads placements:** Placing ads in user-generated content
 - Place an ad when one praises an product
 - Place an ad from a competitor if one criticizes a product

Affect (psychology)

- "Affect" can mean an instinctual reaction to stimulation that occurs before the typical cognitive processes considered necessary for the formation of a more complex emotion.
- This reaction to stimuli is primary for human beings and that it is the dominant reaction for non-human organisms.
- Affective reactions can occur without extensive perceptual and cognitive encoding and be made sooner and with greater confidence than cognitive judgments

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*

What is Sentiment Analysis

Sentiment analysis is the detection of **attitudes**

“enduring, affectively colored beliefs, dispositions towards objects or persons”

1. **Holder (source)** of attitude
2. **Target (aspect)** of attitude
3. **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

4. **Text** containing the attitude

Sentence or entire document

Sentiment Analysis

- Simplest task:
 - Is the attitude of this text positive or negative?
- More complex:
 - Rank the attitude of this text from 1 to 5
- Advanced:
 - Detect the target, source, or complex attitude types

Opinion Mining Using TextBlob

- Tokenize the tweet
- Remove stop words from the tweet, which are not relevant for the mining.
- Parts of Speech Tagging is the most important activity done by TextBlob object. Here, adjectives, adverbs etc. which are relevant are tagged to the keyword.
- Passing the processed tokens to Sentiment Classifier which will return a value between -1.0 and 1.0

How the Sentiment Classifier works

- TextBlob uses a large Movie Review Dataset which is pre-classified as positive and negative.
- TextBlob trains using the Naive Bayes classifier to determine positive and negative reviews.
- Now, the sentiment classifier essentially calculates the polarity of tokens between -1.0 and 1.0
- -1.0 is negative, 0.0 is neutral and 1.0 is positive

So, what is Subjectivity?

- The linguistic expression of somebody's opinions, sentiments, emotions.....

Tweepy Code

- `tweets = api.user_timeline(screen_name="realDonaldTrump", count=200)`
- `print("5 recent tweets:\n")`
- `for tweet in tweets[:5]:`
- `print(tweet.text)`

Sentiment analysis - Levels

- document level,
- sentence level or
- aspect/feature level.

Document Level Classification

- In this process, sentiment is extracted from the entire review, and a whole opinion is classified based on the overall sentiment of the opinion holder.
- Goal is to classify a review as positive, negative, or neutral.
- Document level classification works best when the document is written by a single person and expresses an opinion/sentiment on a single entity.

Sentence Level Classification – 2 Steps

- Subjectivity classification of a sentence into one of two classes:
 - objective and subjective
- Sentiment classification of subjective sentences into two classes:
 - positive and negative

Sentence Level Classification

- An objective sentence presents some factual information
- Subjective sentence expresses personal feelings, views, emotions, or beliefs.
- Subjective sentence identification can be achieved through different methods such as Naïve Bayesian classification.
- However, just knowing that sentences have a positive or negative opinion is not sufficient.
- This is an intermediate step that helps filter out sentences with no opinions and helps determine to an extent if sentiments about entities and their aspects are positive or negative.
- ~~• A subjective sentence may contain multiple opinions and subjective and factual clauses.~~

Aspect/Feature Level Classification

- In this process, the goal is to identify and extract object features that have been commented on by the opinion holder and determine whether the opinion is positive, negative, or neutral.
- Feature synonyms are grouped, and a feature-based summary of multiple reviews is produced.

