Natural Language Understanding, Generation, and Machine Translation

Lecture 28: Sentiment Analysis

Laura Perez-Beltrachini

(Slides by Mirella Lapata)

School of Informatics
University of Edinburgh
mlap@inf.ed.ac.uk

Outline

What is Sentiment Analysis?

STAR WARS: THE RISE OF SKYWALKER

Critics Consensus

Star Wars: The Rise of Skywalker suffers from a frustrating lack of imagination, but concludes this beloved saga with fan-focused devotion.

52% **86%**

TOMATOMETER
Total Count: 471

AUDIENCE SCORE
Verified Ratings: 93,869

MORE INFO

What is Sentiment Analysis?

January 8, 2020

Star Wars: The Rise of Skywalker is a good tentpole movie for the holiday season and a fun Star Wars movie that ends on a high.

Full Review | Original Score: 8/10

January 9, 2020

Kudos to J.J. Abrams for doing something extraordinary: he has made me not care about Star Wars for the first time ever. I'm kind of relieved that it's over, because it has stopped being fun.

Full Review | Original Score: 2/5

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

1 star	2	3	4 stars	5 stars

What people are saying

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

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 ★★★½☆ (144)

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

Most mentioned

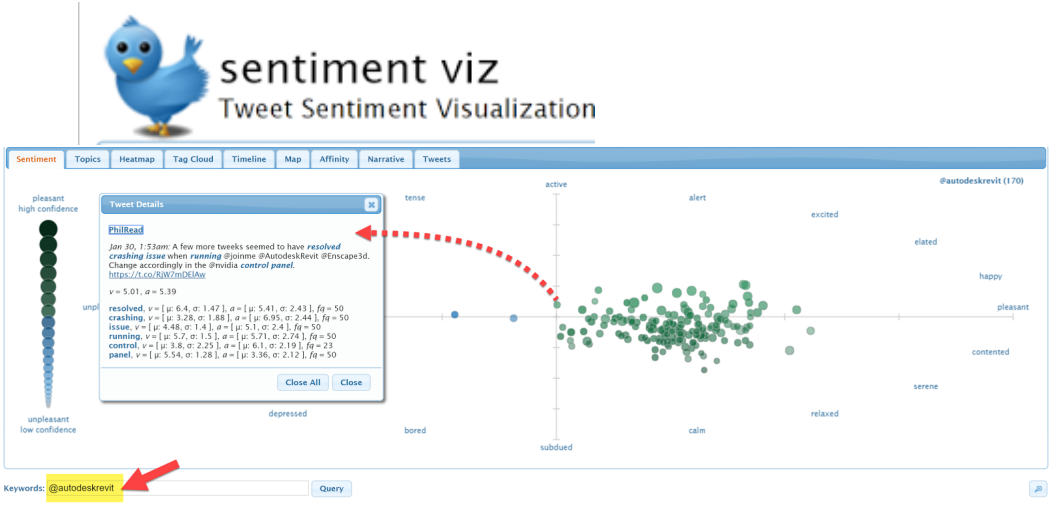
Performance	(57)
Ease of Use	(43)
Print Speed	(39)
Connectivity	(31)

More ▼

Show reviews by source

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

Predicting Sentiment on Twitter



Why Sentiment Analysis?

- **Movies**: 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

Sentiment analysis has many other names: opinion extraction, opinion mining, sentiment mining, subjectivity analysis.

A Typology of Affective States

Scherer's (1984) typology of affective states provides a broad framework for understanding sentiment.

Emotion evaluation of a major event: angry, sad, joyful, fearful, ashamed, proud, elated
Mood diffuse affect state, 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 affectively colored beliefs, dispositions towards objects or persons: liking, loving, hating, valuing, desiring
Personality traits personality dispositions and typical behavior tendencies: nervous, anxious, reckless, morose, hostile, jealous

A Typology of Affective States

Scherer's (1984) typology of affective states provides a broad framework for understanding sentiment.

Emotion evaluation of a major event: angry, sad, joyful, fearful, ashamed, proud, elated
Mood diffuse affect state, 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 affectively colored beliefs, dispositions towards objects or persons: liking, loving, hating, valuing, desiring
Personality traits personality dispositions and typical behavior tendencies: nervous, anxious, reckless, morose, hostile, jealous

Sentiment Analysis

Sentiment analysis is the detection of attitudes:

- **Holder (source)** of attitude
- **Target** of attitude
- **Type** of attitude (like, love, hate, value desire)
- More commonly simple **polarity**: positive, negative, neutral
- **Text** containing the attitude: words, sentences, documents

Online User Reviews

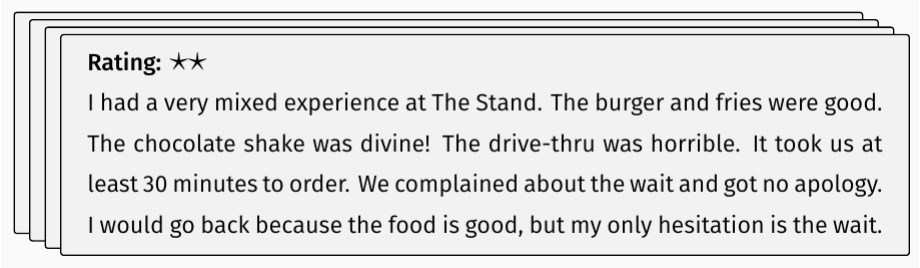
Heavily influence customer decisions:

- Travel booking (Ye et al., 2009)
- Box Office success (Duan et al., 2008)
- Shopping (TurnTo.com report, 2018)

Incredibly rich data source:

- 6.3 million Yelp reviews written in 2010
- 27.3 million in 2017

Document-level Sentiment Analysis



[insert favourite neural net here]

[Johnson and Zhang (2015); Yang et al. (2016); Liu and Lapata (2018)]

⇓

Predicted rating: ⋆⋆

Fine-grained Sentiment Analysis

Rating: ★★

I had a very mixed experience at The Stand. The burger and fries were good. The chocolate shake was divine! The drive-thru was horrible. It took us at least 30 minutes to order. We complained about the wait and got no apology. I would go back because the food is good, but my only hesitation is the wait.

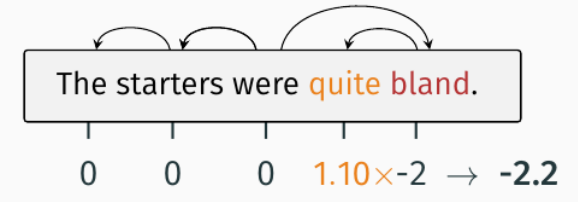
Positive:

- The burger and fries were good.
- The chocolate shake was divine!
- I would go back because the food is good.

Negative:

- The drive-thru was horrible.
- It took us at least 30 minutes to order.
- We complained about the wait and got no apology.
- My only hesitation is the wait.

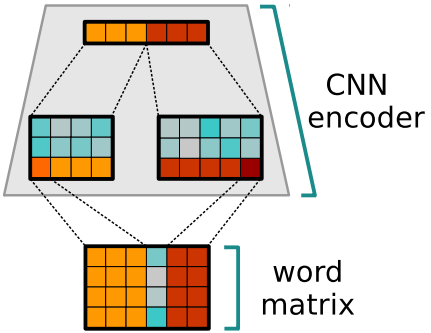
Fine-grained Sentiment: Unsupervised Methods



Adjective:		**Intensifier:**	
disgusting	-5	slightly	0.50
terrible	-4	somewhat	0.70
bland	-2	pretty	0.90
so-so	-1	quite	1.10
okay	1	really	1.15
great	2	very	1.25
amazing	4	extraordinarily	1.50
divine	5	(the) most	2.00

SO-CAL: Semantic Orientation CALculator (Taboada et al., 2011)

Fine-grained Sentiment: Supervised Methods



- Segment CNN (Kim, 2014)
- multiple convolutional filters of varying length
- max-over-time pooling
- **Successful for sentence classification** ☺
- **Segment encoder in larger networks** ☺
- **Requires expensive annotations** ☹

Fine-grained Sentiment: Weakly-supervised Methods

Modeling Assumptions:

- Build a document-level sentiment classifier
- We can obtain reviews (documents) and labels (ratings) for free!
- Hierarchical model: first build representations of sentences, then aggergate those into a document vector (Tang et al., 2015)
- Sentences contribute differentially to document representations (Yang et al., 2016)
- Use document-level classifier to label segments.

Fine-grained Sentiment: Hierarchical Network (HIERNET)

Input: Word Matrices $\mathbf{X}_i$

Segment encoding:

$$\mathbf{v}_i = \mathrm{CNN}(\mathbf{X}_i)$$

Document encoding:

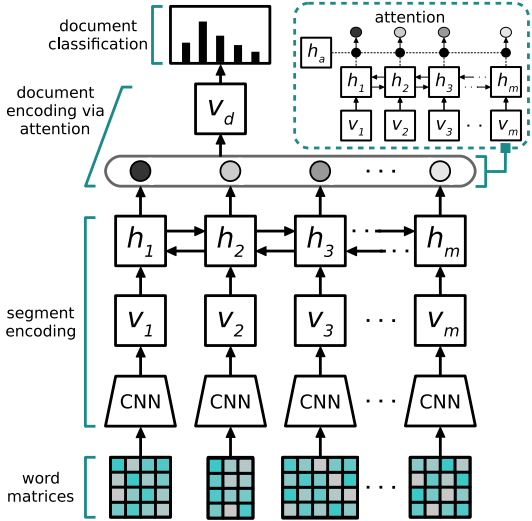
$$\overrightarrow{\mathbf{h}}_i = \overrightarrow{\mathrm{GRU}}(\mathbf{v}_i),\ \overleftarrow{\mathbf{h}}_i = \overleftarrow{\mathrm{GRU}}(\mathbf{v}_i),$$
$$\mathbf{h}_i = [\overrightarrow{\mathbf{h}}_i, \overleftarrow{\mathbf{h}}_i],\ i \in [1, m]\,.$$

Document classification:

$$\mathbf{p}_r = \mathrm{softmax}(\mathbf{W}_c \mathbf{v}_i + \mathbf{b}_c)$$

Objective:

NLL of document predictions



Fine-grained Sentiment: A Note on Attention

Simple averaging produces document-level representation: $\mathbf{v}_R = \frac{1}{m}\sum_y \mathbf{h_i}$

Importance of each segment is measure via vector $\mathbf{h}_a$ (trained key, able to recognize sentiment-heavy segments):

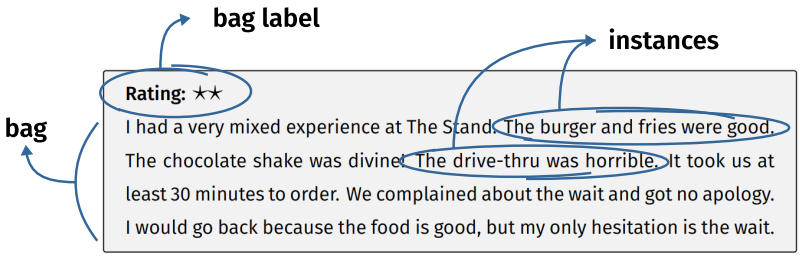
$$\mathbf{h}'_i = \tanh(\mathbf{W}_a\mathbf{h}_i + \mathbf{b}_a)$$

$$a_i = \frac{exp(\mathbf{h}_i'^{\mathrm{T}}\mathbf{h}_a)}{\sum_i exp(\mathbf{h}_i'^{\mathrm{T}}\mathbf{h}_a)}$$

Fine-grained Sentiment: Weakly-supervised Methods

Multiple Instance Learning (MIL; Keeler and Rumelhart, 1992)*

- Training examples (reviews) → bags of instances
- Bag labels (ratings) → supervision
- Instance labels → latent
- Bag-instance relationship?



* **MIL In NLP:** Surdeanu et al. (2012); Kotzias et al. (2015); Pappas and Popescu-Belis (2017)

Fine-grained Sentiment: Weakly-supervised Methods

Modeling Assumptions:

- Segment s_i conveys sentiment polarity: $pol_i \in [-1, +1]$
- Segments have varying degrees of importance: $a_i \in [0, 1]$, $\sum_i a_i = 1$
- Overall polarity of review: average of polarities, weighted by importance
- Segments can be words, phrases, clause, sentences

Multiple Instance Learning Network (MILNET)

Input: Word Matrices $\mathbf{X}_i$

Segment encoding:

$\mathbf{v}_i = \mathrm{CNN}(\mathbf{X}_i)$

Segment classification:

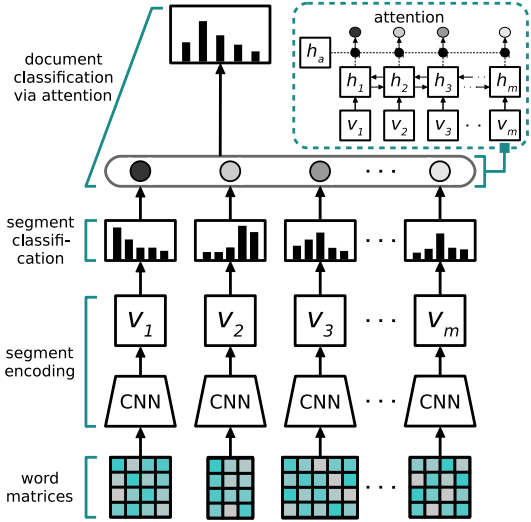
$\mathbf{p}_i = \mathrm{softmax}(\mathbf{W}_c \mathbf{v}_i + \mathbf{b}_c)$

Document classification:

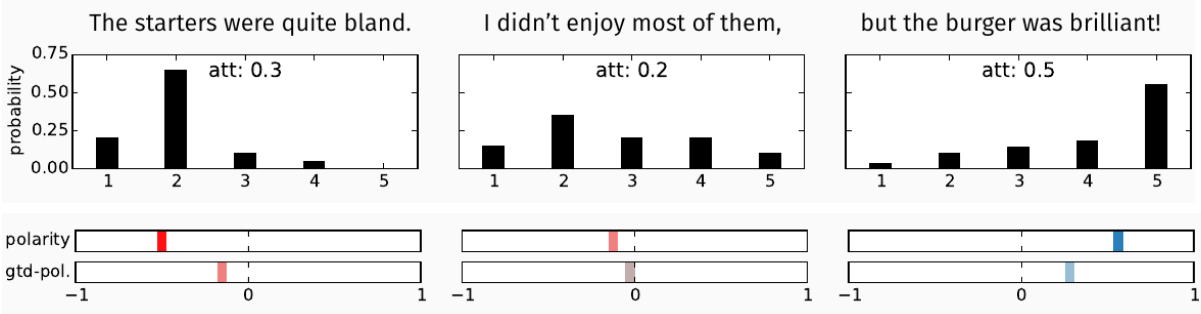
$p_r^{(c)} = \sum_i a_i p_i^{(c)}, c \in \{1, C\}$

Objective:

NLL of document predictions



Polarity Scoring



Polarity of segment: $pol_i = \sum_c p_i^{(c)} w^{(c)}$, $\mathbf{w} = \langle -1, -0.5, 0, +0.5, +1 \rangle$

Gated polarity $\rightarrow$ accounts for segment importance: $gpol_i = a_i \cdot pol_i$

Polarity-based Opinion Extraction

Rating: ⋆⋆

I had a very mixed experience at The Stand. The burger and fries were good. The chocolate shake was divine! The drive-thru was horrible. It took us at least 30 minutes to order. We complained about the wait and got no apology. I would go back because the food is good, but my only hesitation is the wait.

Very negative ↔ Very positive

[+1.00]	**The chocolate shake was divine**
[+0.86]	**I would go back because the food is good**
[+0.50]	**The burger and fries were good**
[-0.05]	I had a very mixed experience at The Stand.
[-0.10]	but my only hesitation is the wait
[-0.10]	and got no apology
[-0.25]	We complained about the wait
[-0.43]	**It took us at least 30 minutes to order**
[-0.89]	**The drive-thru was horrible**

Discrete predictions

⇓

Gated polarities

⇓

Rankings

⇓

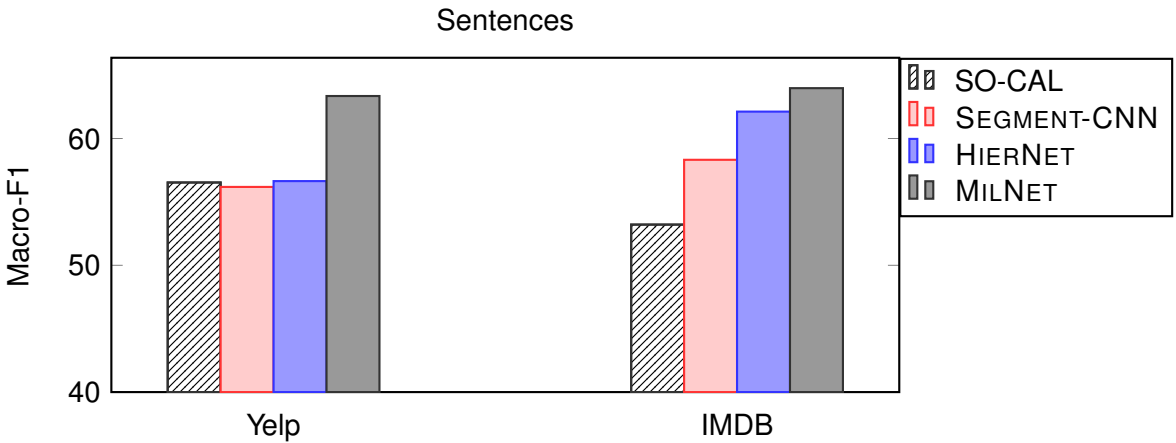
Opinion extraction

Experimental Setup: Datasets

Document-level	**Yelp'13**	**IMDB**
Documents	335K	348K
Avg # Sentences	8.90	14.02
Avg # Words	152	325
Vocabulary Size	129K	97K
Classes	1–5	1–10

Segment-level	**Yelp'13**	**IMDB**
Documents	100	100
Sentences	1,065	1,029
Classes	{– , 0 , +}	

Results


Sentences
Macro-F1
60
50
40
Yelp
IMDB
SO-CAL
SEGMENT-CNN
HIERNET
MILNET

Discussion

- A MIL neural model for fine-grained sentiment analysis
- Attention-based polarity scoring method facilitates opinion extraction
- Model is trained on document-level labels only!

The good things are the acting. Mostly brilliant, and believable. On the negative side is, well everything else. I bet even the catering was bad on this film.

Very tasty and fresh, I really enjoyed it. Our server was a bit aloof! Very sweet girl though. Haha!

I would give zero stars. it was ice cold. This was torture! The staff is clueless. Horrible service!