Sentiment Analysis

Web Science Lecture

1 March 2020

Sagnik Ray Choudhury
src@di.ku.dk
http://sagnik.in

UNIVERSITY OF COPENHAGEN

Organization

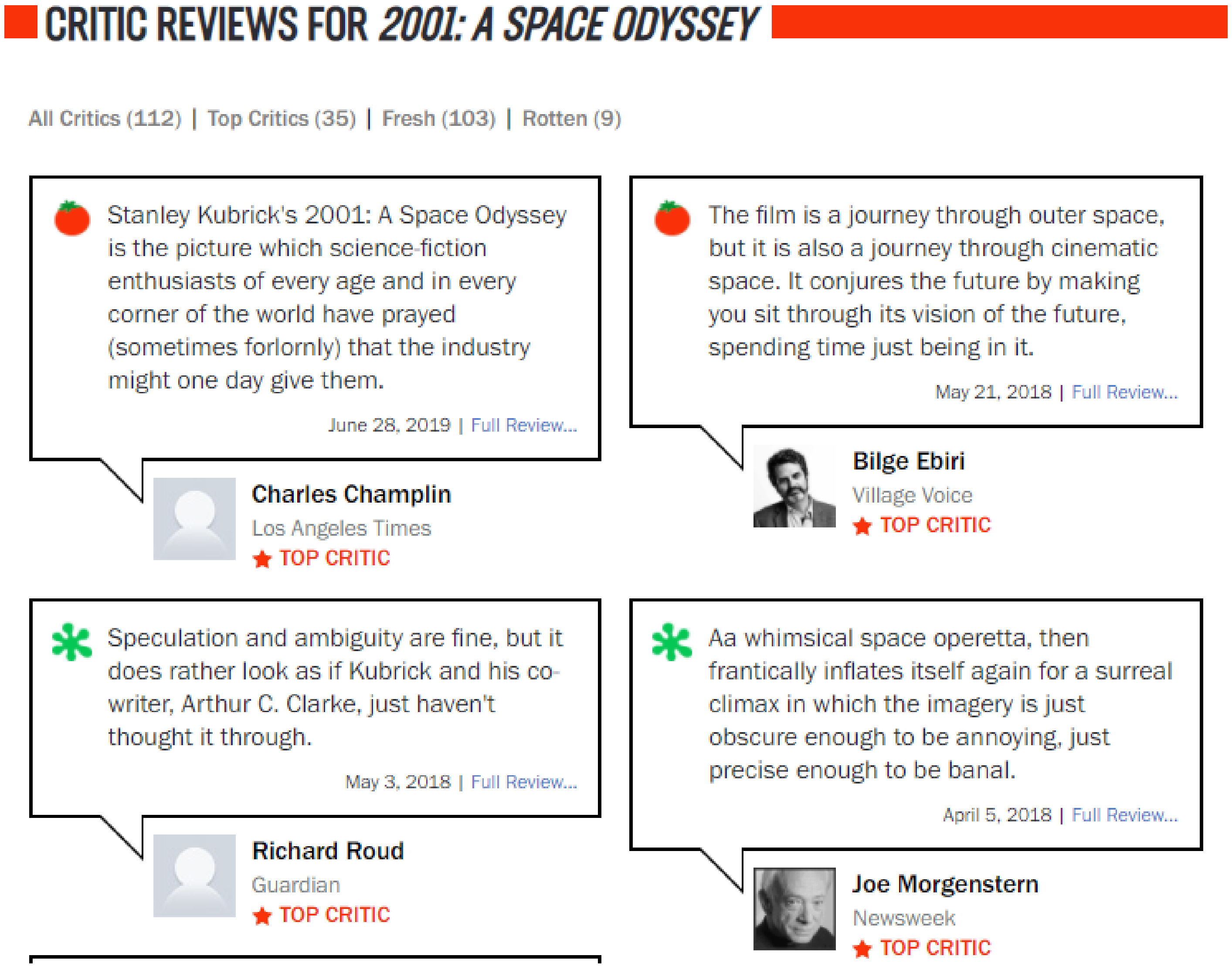
- Part 1: problem definition, applications
- Part2: unsupervised methods, sentiment lexica
- Part3: supervised methods for sentiment analysis
  - We will see one representation of a sentence, the other representations in the next lecture.
- Part4: research areas
- Acknowledgements:
  - Isabelle Augenstein (5100-B3-3F20;Web Science)
  - Christopher Potts (Stanford CS224U)
  - Christopher Manning (Stanford CS224N)
  - Internet

Problem Definition, Variations

- Sentiment Analysis
  - Automatically identify the feeling expressed in a text
  - Detect words/phrases that express such feelings
  - "I am *concerned* that the coronavirus outbreak is already *hurting the world economy*"
- Two types of sentiment analysis
  - Target based: sentiment towards a concrete target, e.g., a product. ``I hate the new iphone 7".
  - Aspect based: Sentiment towards a concrete sub-aspect of a target (price, cleanliness, size): "I dislike the *camera* of iPhone 7" (also, opinion mining, Dave et.al. 2003).
  - See details in section 1.5 of Pang and Lee, 2008.
- Stance detection: similar to aspect-based sentiment analysis, but the target is not mentioned.

No more #NastyWomen or #BadHombres

Application 1: Movie Review: Rotten Tomatoes



CRITIC REVIEWS FOR *2001: A SPACE ODYSSEY*

All Critics (112) | Top Critics (35) | Fresh (103) | Rotten (9)

Stanley Kubrick's 2001: A Space Odyssey is the picture which science-fiction enthusiasts of every age and in every corner of the world have prayed (sometimes forlornly) that the industry might one day give them.

June 28, 2019 | Full Review...

Charles Champlin
Los Angeles Times
TOP CRITIC

The film is a journey through outer space, but it is also a journey through cinematic space. It conjures the future by making you sit through its vision of the future, spending time just being in it.

May 21, 2018 | Full Review...

Bilge Ebiri
Village Voice
TOP CRITIC

Speculation and ambiguity are fine, but it does rather look as if Kubrick and his co-writer, Arthur C. Clarke, just haven't thought it through.

May 3, 2018 | Full Review...

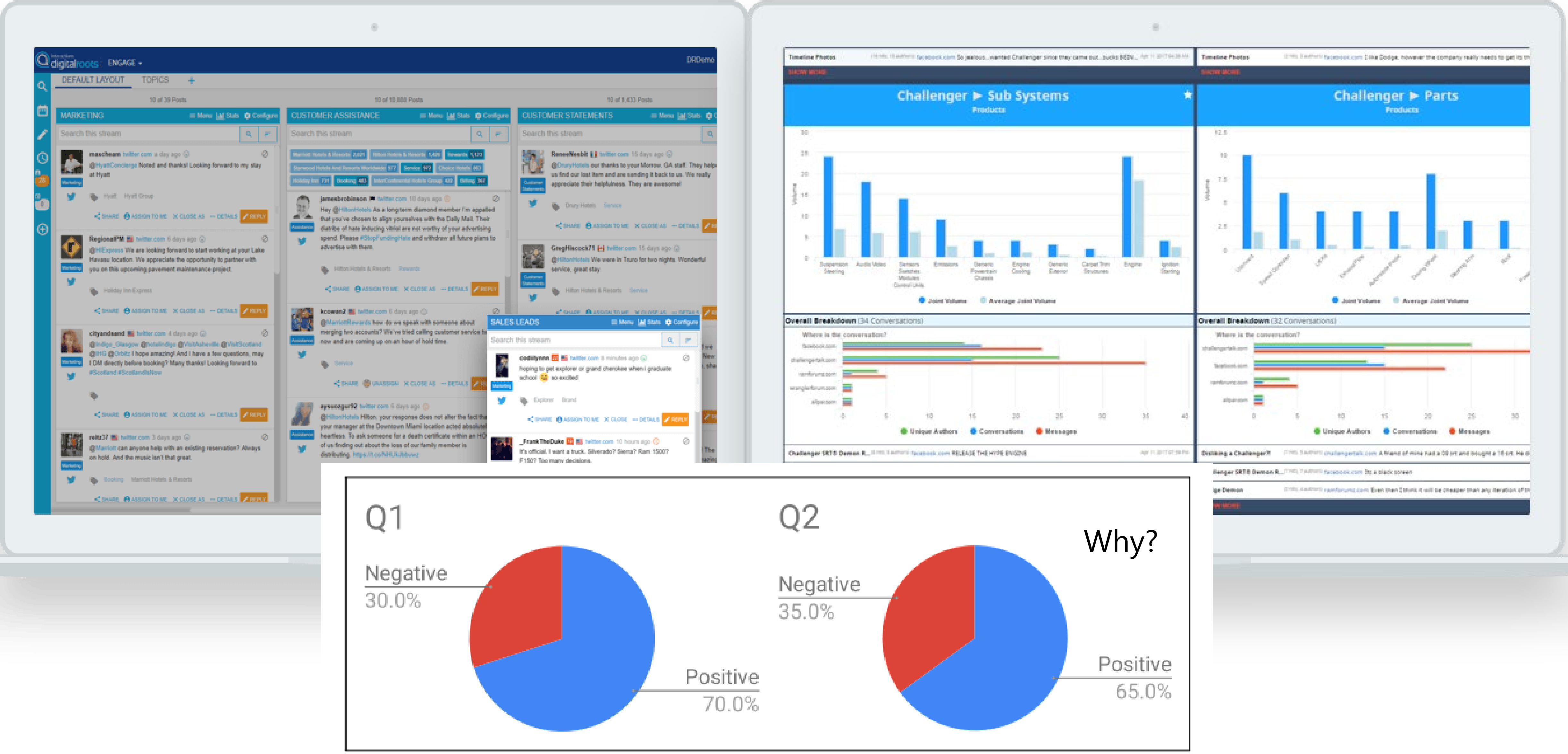
Richard Roud
Guardian
TOP CRITIC

Aa whimsical space operetta, then frantically inflates itself again for a surreal climax in which the imagery is just obscure enough to be annoying, just precise enough to be banal.

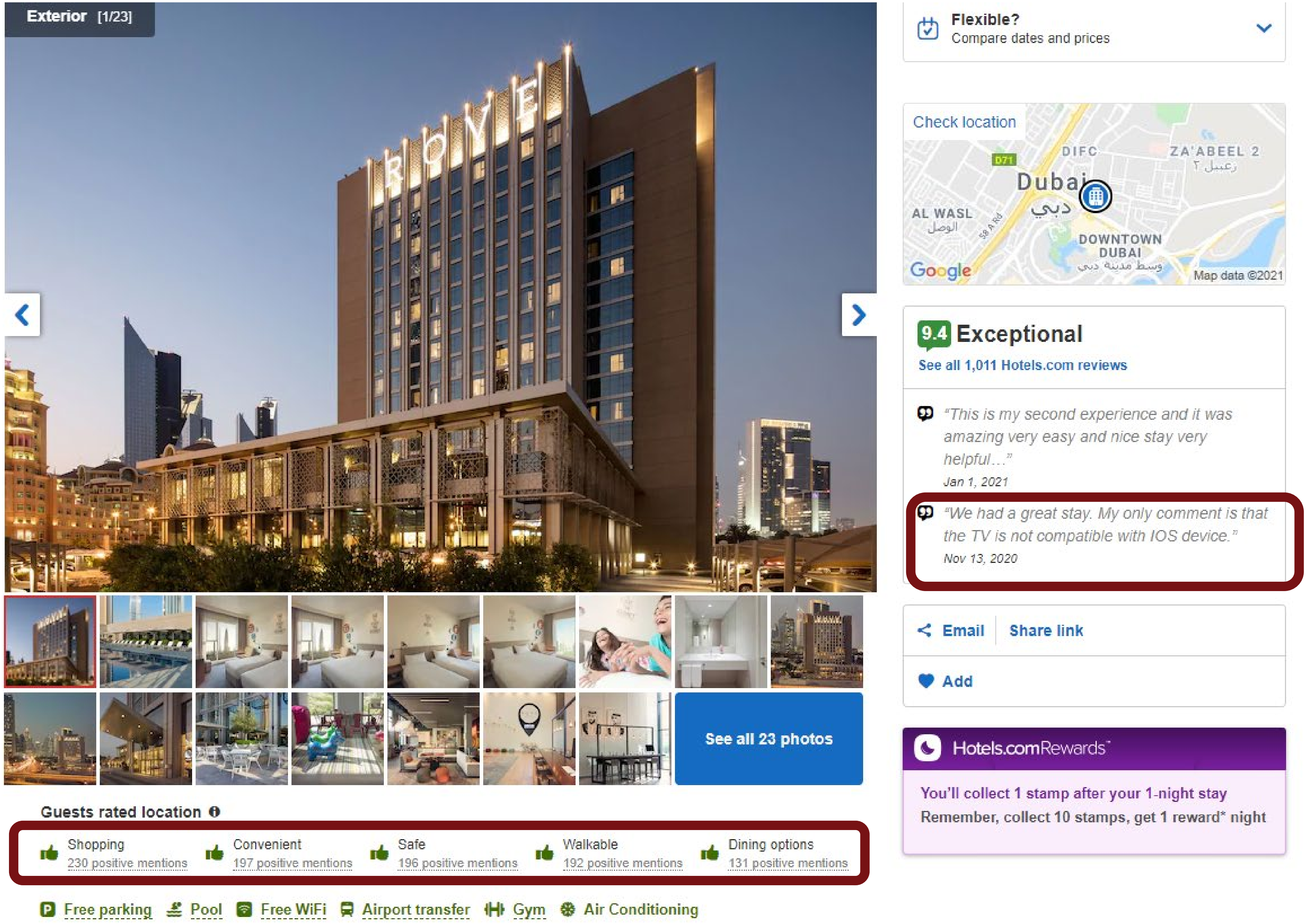
April 5, 2018 | Full Review...

Joe Morgenstern
Newsweek
TOP CRITIC

Application 2: B2B Dashboards



Application 3: Opinion Mining on Hotels.com



Both positive and negative sentiment

Different aspects

Application 4: Story Arcs in Novels

Reagan, 2016

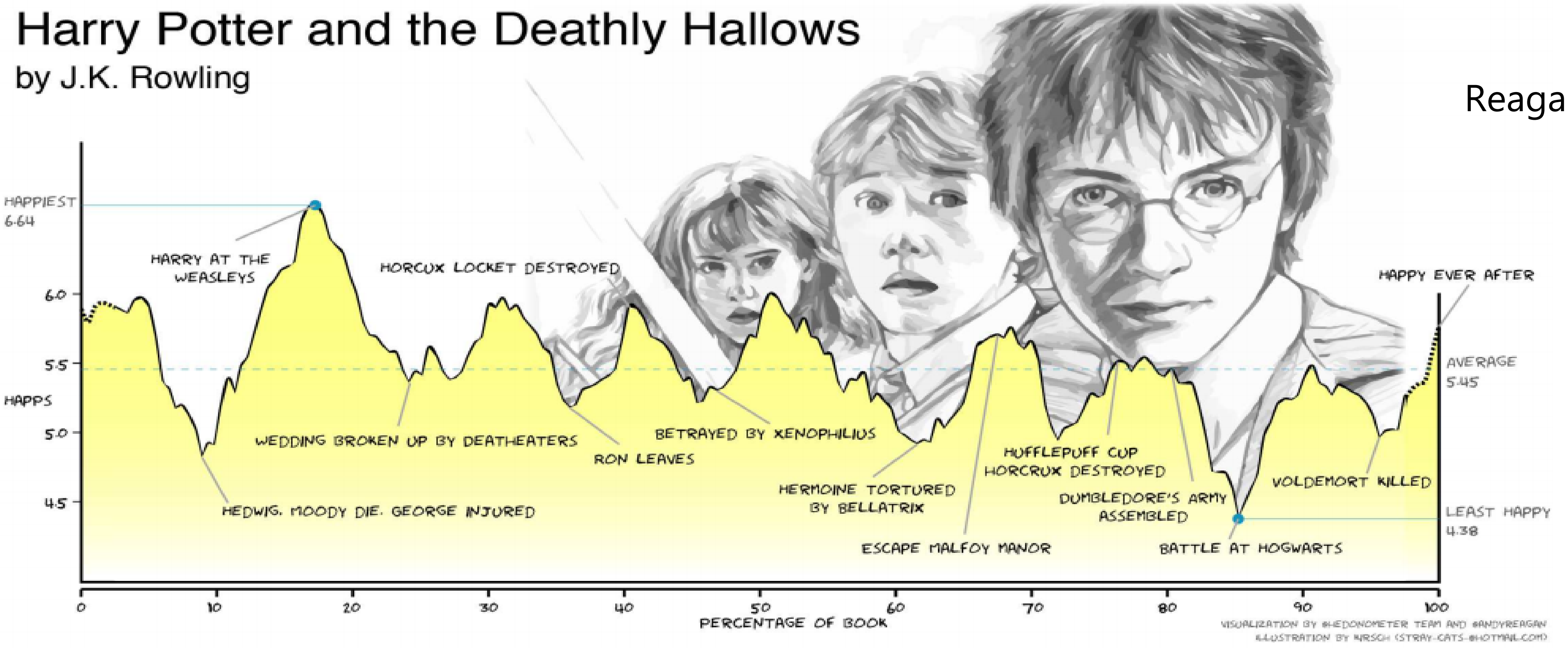


FIG. 2: Annotated emotional arc of *Harry Potter and the Deathly Hallows*, by J.K. Rowling, inspired by the illustration made by Medaris for The Why Files [23]. The entire seven book series can be classified as a "Kill the monster" plot [24], while the many sub plots and connections between them complicate the emotional arc of each individual book: this plot could not be readily inferred from the emotional arc alone. The emotional arc shown here, captures the major highs and lows of the story, and should be familiar to any reader well acquainted with Harry Potter. Our method does not pick up emotional moments discussed briefly, perhaps in one paragraph or sentence (e.g., the first kiss of Harry and Ginny). We provide interactive visualizations of all Project Gutenberg books at `http://hedonometer.org/books/v3/1/` and a selection of classic and popular books at `http://hedonometer.org/books/v1/`.

How Subjective is Sentiment?

- Can we identify the sentiment (positive, negative, neutral) in the following sentences?
  - There was an earthquake in California.
  - They said it would be great.
  - They said it would be great, and they were right.
  - They said it would be great, and they were wrong.
  - Oh, you’re terrible!

Summary

- What are the different dimensions of sentiment analysis?
- Why is the problem important?
- What makes the problem difficult? : subjectivity

Sentiment Analysis: Task

- Does a piece of text express a positive, negative or neutral sentiment?

Cathy Polinsky
@cathy_polinsky
Follow

So proud of @stitchfix technology team for increasing our gender diversity from 31% to over 35% women in the past year! Happy IWD!

Bajas K. Smith
@jdnaa
Follow

Replying to @cathmckenna

Oh yay. International Womans Day. Just another day for the government to waste tax payers money on a day that means nothing for most women, myself included. When is International Mens Day?

The Economist
@TheEconomist
Following

Jetex has opened 39 private-jet terminals in more than 20 countries since 2005

Large Datasets for Fine Grained Sentiment Analysis

- IMDb movie reviews (50K) (Maas et al. 2011): http://ai.stanford.edu/~amaas/data/sentiment/index.html
- Datasets from Lillian Lee's group: http://www.cs.cornell.edu/home/llee/data/
- Datasets from Bing Liu's group: https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html
- RateBeer (McAuley et al. 2012; McAuley and Leskovec 2013): http://snap.stanford.edu/data/web-RateBeer.html (do you like beer?)
- Amazon Customer Review data: https://s3.amazonaws.com/amazon-reviews-pds/readme.html

(very large, kind of hard to access)

- Amazon Product Data (McAuley et al. 2015; He and McAuley 2016): http://jmcauley.ucsd.edu/data/amazon/ (amazing because subsets of data are available for experiments)
- Stanford Sentiment Treebank (SST; Socher et al. 2013) https://nlp.stanford.edu/sentiment/
- http://nlpprogress.com/english/sentiment_analysis.html

Unsupervised Methods for Sentiment Analysis

- How would you do it?
  - Construct gazetteers(lists of) of positive and negative terms.
  - Count positive and negative terms in the text.
  - **Positive**: admire, amazing, assure, celebration, charm, eager, enthusiastic, excellent, fancy, fantastic, frolic, graceful, happy, joy ···
  - **Negative**: abominable, anger, anxious, bad, catastrophe, cheap, complaint, condescending, deceit, defective ···
- "I had an excellent time celebrating my birthday at the wonderful hotel X" -> positive
- "I had an excellent time despite the unfriendly waiting staff" -> neutral

Existing Sentiment Lexica

- What would it look like?
  - positive: List[Words], negative: List[Words], neutral: List[Words]
  - word: positive score, negative score.
  - **SentiWordNet**: https://github.com/aesuli/SentiWordNet

```
a	00005107	0.5	0	uncut#7 full-length#2	complete; "the full-length play"
a	00005205	0.5	0	absolute#1	perfect or complete or pure; "absolute loyalty"; "absolute silence"; "absolute truth"; "absolute alcohol"
a	00005473	0.75	0	direct#10	lacking compromising or mitigating elements; exact; "the direct opposite"
a	00005599	0.5	0.5	unquestioning#2 implicit#2	being without doubt or reserve; "implicit trust"
a	00005718	0.125	0	infinite#4	total and all-embracing; "God's infinite wisdom"
a	00005839	0.5	0.125	living#3	(informal) absolute; "she is a living doll"; "scared the living daylights out of them"; "beat the living hell out of him"
a	00006032	0.25	0.5	relative#1 comparative#2	estimated by comparison; not absolute or complete; "a relative stranger"
a	00006245	0	0	relational#1	having a relation or being related
a	00006336	0	0	absorptive#1 absorbent#1	having power or capacity or tendency to absorb or soak up something (liquids or energy etc.); "as absorbent as a sponge"
a	00006777	0.375	0	sorbefacient#1 absorbefacient#1 inducing or promoting absorption
```

The columns are: 1. part of speech tag. 2. id, 3. PosScore, 4. negScore, 5. Synset terms, 6. Gloss

Why might we need a part of speech tag?

Word	PoS Tag	Sentiment
Fine	Adj	+ve
Fine	Verb	-ve
Pass	Noun	+ve
Pass	Verb	-ve

Other Sentiment Lexica

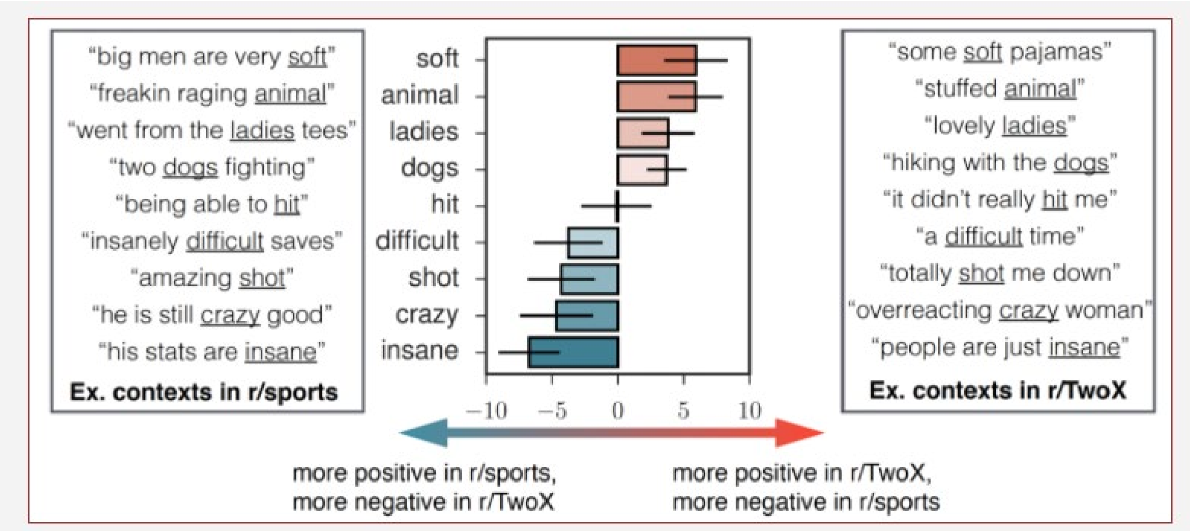
- Bing Liu's Opinion Lexicon: nltk.corpus.opinion_lexicon
- SentiWordNet: nltk.corpus.sentiwordnet
- MPQA subjectivity lexicon: http://mpqa.cs.pitt.edu (phrases and contexts)
- Linguistic Inquiry and Word Counts (LIWC): https://liwc.wpengine.com
- Hamilton et al. (2016): SocialSent https://nlp.stanford.edu/projects/socialsent/ (reddit communities, also historical trends)
- Brysbaert et al. (2014): Norms of valence, arousal, and dominance for 13,915 English lemmas.
- WordNet Affect.

(5) Thousands of coup supporters celebrated (**positive**) overnight, waving flags, blowing whistles ...

(6) The criteria set by Rice are the following: the three countries in question are repressive (**negative**) and grave human rights violators (**negative**) ...

(7) Besides, politicians refer to good and evil (**both**) only for purposes of intimidation and exaggeration.

(8) Jerome says the hospital feels (**neutral**) no different than a hospital in the states.



Disagreement Levels in Sentiment Lexica

- Can we combine different sentiment lexica?
  - Yes, but keep in mind they are not intended to capture the same *aspect* of sentiment
- Are there disagreements?

	MPQA	Opinion lexicon	Inquirer	SentiWordNet	LIWC
MPQA	-	33/5402 (0.6%)	49/2867 (2%)	1127/4214 (2%)	12/363 (3%)
Opinion Lexicon	-	-	32/2411 (1%)	1004/3994 (25%	9/403(2%)
Inquirer	-	-	-	520/2306 (23%)	1/204 (0.5%)
SentiWordNet	-	-	-	-	174/694 (25%)
LIWC	-	-	-	-	-

- Why do we even need sentiment lexica, are most things not supervised using DNN anyway?
  - Domain dependence and getting domain specific data is hard.
  - ``fast car" vs ``that was over fast"
  - UCPH is [MASK] -> university (neutral)
  - Breaking Bad is [MASK] -> awesome (positive)

Semantic Orientation of Phrases

- Sentiment classification of product reviews: recommended or not recommended
- Step 1: Extract phrases (bigrams) with some patterns (mostly phrases that contain an adjectives or adverbs)
- Step 2 & 3: Find *semantic orientation* of each phrase (positive and negative) and average.

$$\mathrm{PMI}(word_1, word_2) = \log_2\left[\frac{\mathrm{p}(word_1 \,\&\, word_2)}{\mathrm{p}(word_1)\,\mathrm{p}(word_2)}\right]$$

$$\mathrm{SO}(phrase) = \mathrm{PMI}(phrase, \text{“excellent”}) - \mathrm{PMI}(phrase, \text{“poor”})$$

- PMI gives an estimate of how much information we get about w1 when we know about w2 (or vice versa).
- How to estimate p(w1) and p(w2)? In a normal corpus, frequency(w1)/number of words in the corpus.
- We can do a little better (since we have access to a search engine):

$$\log_2\left[\frac{\mathrm{hits}(phrase\ \mathrm{NEAR}\ \text{“excellent”})\ \mathrm{hits}(\text{“poor”})}{\mathrm{hits}(phrase\ \mathrm{NEAR}\ \text{“poor”})\ \mathrm{hits}(\text{“excellent”})}\right]$$

- Understanding information theory: Stanford course project.

Summary

- Different sentiment datasets
- Different sentiment lexica
  - Are sentiment lexica relevant in the era of deep learning?
- Unsupervised methods of sentiment classification

Supervised Sentiment Classification

- Basic setup is pretty simple: f(sentence) -> positive, negative, neutral
- Two things to consider:
  - How to represent a sentence? IOW, sentence -> vector
  - How to *approximate* f?
- How to represent a sentence?
  - The basic unit of sentence (tokens)
  - Why tokens and not “words”?
  - We can represent a text by the words in it:
  - Consider your vocabulary has these words {‘no’, ‘more’, ‘nasty’, ‘woman’, ‘or’, ‘bad’, ‘hombres’, ‘say’, ‘yes’, ‘to’, ‘love’}
  - You can represent the sentence as

no	more	nasty	woman	or	bad	hombres	say	yes	to	love
1	1	1	1	1	1	1	0	0	0	0

- This is called the vector space model (the document is represented in the vector space of terms).
- You can use tf-idf (term frequency and inverse document frequency) instead of binary values.

Sentence Representation: Features

- How to represent a sentence with “features”?
- What can be a feature:
  - Part of speech tags.
  - Use the sentence lexica as the vocabulary.

	(like, V)	(love, V)	…	(hate, V)	(bad, ADJ)
I like the movie	1	0	…	0	0
I hate the game	0	0	…	1	0

- Features are computed on tokens: very sparse representation.
- Features can be arbitrary: is there an adjective after a verb?
  - What other features can you think of?
- Some practical considerations
  - Tokenizers
  - Stemmers

Practical Considerations: Tokenizer

- "Mr. President went to Washington D.C., but the first lady stayed back."
- Tokenizer should:
  - Isolate emoticons
  - Respects Twitter and other domain-specific markup
  - Uses the underlying mark-up (e.g., <strong> tags)
  - Preserves capitalization
  - Regularizes lengthening (e.g., YAAAAAAY⇒YAAAY)
  - Captures significant multiword expressions (e.g., out of this world)
- Whitespace tokenizer, treebank tokenizer, nltk.tokenize.casual.TweetTokenizer

Practical Considerations: Stemmer

- Stemming and lemmatization: Reduce inflectional forms and sometimes derivationally related forms of a word to a common base form.
  - Stemming: a crude heuristic process that chops off the ends of words.
  - Lemmatization : 1. remove inflectional endings only, 2. To return the base or dictionary form of a word.

    am, are, is $\Rightarrow$ be

    car, cars, car's, cars' $\Rightarrow$ car

Stemming, Manning IR book

- Common English Stemmers:
  - The Porter stemmer
  - The Lancaster stemmer
  - The WordNet stemmer

Stemmer problems

Porter

Positiv	Negativ	Porter stemmed
defense	defensive	defens
extravagance	extravagant	extravag
affection	affectation	affect
competence	compete	compet
impetus	impetuous	impetu
objective	objection	object
temperance	temper	temper
tolerant	tolerable	toler

Lancaster

Positiv	Negativ	Lancaster stemmed
call	callous	cal
compliment	complicate	comply
dependability	dependent	depend
famous	famished	fam
fill	filth	fil
flourish	floor	flo
notoriety	notorious	not
passionate	passe	pass
savings	savage	sav
truth	truant	tru

Wordnet

Positiv	WordNet stemmed
(exclaims, v)	exclaim
(exclaimed, v)	exclaim
(exclaiming, v)	exclaim
(exclamation, n)	exclamation
(proved, v)	prove
(proven, v)	prove
(proven, a)	proven
(happy, a)	happy
(happier, a)	happy
(happiest, a)	happy

- Is stemming/lemmatization a bad practice?
- Why did we start doing them?

Practical Considerations: Negation

- Negation makes sentiment classification difficult.
  - I liked the movie.
  - I didn’t like the movie.
  - I don’t like it, but someone else might.
  - I don’t think I will like the movie, but
- A very simple improvisation
  - Append a _NEG suffix to the words after a negation and a clause delimiter.
  - How do we detect negation? (dictionary based)

I	don’t	like	it	but	someone	else	might
I	don’t	Like_NEG	It_NEG	but	someone	else	might

Word Representations: Primer

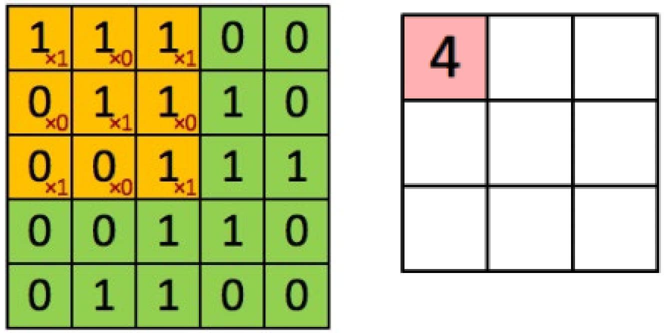
- So far in sentiment analysis:
  - Unsupervised, supervised methods.
  - Representation of sentences in vector space model (words are the features).
  - Representation of sentences by other features (sentiment lexica, arbitrary features)
- But what if we could put the words in a “semantic” space? In other words, have a continuous vector representation of the words so that we can say
  - love and like are synonyms.
  - love and hate are antonyms.
- “You shall know a word by the company it keeps” (Firth, 1957).
- We will discuss this in detail in the next week.
- For now, let’s assume that we have a fixed length vector for each word in our vocabulary.

Simple Classifiers Using these Representations

- We have fixed length vectors for each word, how do we get a sentence representation (a fixed length vector for each sentence)?
  - mean pooling (a simple average of all vectors)
  - max pooling (max of all vectors -> reduces a sentence to a single word)
- Building sentence representations from words through convolutions.
- What is the idea behind convolution in NLP?
  - Compute a representation (vector) for every possible word subsequence of a certain length.
  - “The Discovery One and its revolutionary super computer seek a mysterious monolith that first appeared at the dawn of man.”:
  - (the discovery one), (discovery one and), (one and its)···
  - Hopefully, a couple of the sequences will capture the essence of the sentence.
  - Problem: What is the proper length of the sequence? How about long range dependencies?

Convolution in NLP: Some Worked Out Examples

- What is a convolution
- 1D convolution for text



w1	movie	0.2	0.1	-0.3	0.4
w2	is	0.5	0.2	-0.3	-0.1
w3	good	-0.1	-0.3	-0.2	0.4
w4	if	0.3	-0.3	0.1	0.1
w5	you	0.2	-0.3	0.4	0.2
w6	are	0.1	0.2	-0.1	-0.1
w7	dumb	-0.4	-0.4	0.2	0.3

3	1	2	-3
-1	2	1	-3
1	1	-1	1

- Kernel of size 3. Why is not the other dimension mentioned?
- stride=1
- Your goal is to learn this kernel.

w1,w2,w3	-1.0
w2,w3,w4	-0.5
w3,w4,w5	-3.6
w4,w5,w6	-0.2
w5,w6,w7	0.3

1 vector for the sentence

Convolution in NLP: Parallel Filters and Padding

- Why might we need padding?
  - Sentence lengths are not the same
- Parallel filters
  - Different views of the data that can be computed in parallel

w0	PAD	0	0	0	0
w1	movie	0.2	0.1	-0.3	0.4
w2	is	0.5	0.2	-0.3	-0.1
w3	good	-0.1	-0.3	-0.2	0.4
w4	if	0.3	-0.3	0.1	0.1
w5	you	0.2	-0.3	0.4	0.2
w6	are	0.1	0.2	-0.1	-0.1
w7	dumb	-0.4	-0.4	0.2	0.3
w8	PAD	0	0	0	0

3	1	2	-3
-1	2	1	-3
1	1	-1	1

1	0	0	1
1	0	-1	-1
0	1	0	1

1	-1	2	-1
1	0	-1	3
0	2	2	1

w0,w1,w2	-0.6	0.2	1.4
w1,w2,w3	-1.0	1.6	-1.0
w2,w3,w4	-0.5	-0.1	0.8
w3,w4,w5	-3.6	0.3	0.3
w4,w5,w6	-0.2	0.1	1.2
w5,w6,w7	0.3	0.6	0.9
w6,w7,w8	-0.5	-0.9	0.1
Max pool over time	0.3	1.6	1.4
Avg pool over time	-0.87	0.26	0.53

Sentence (sentiment) Classification Using Convolution

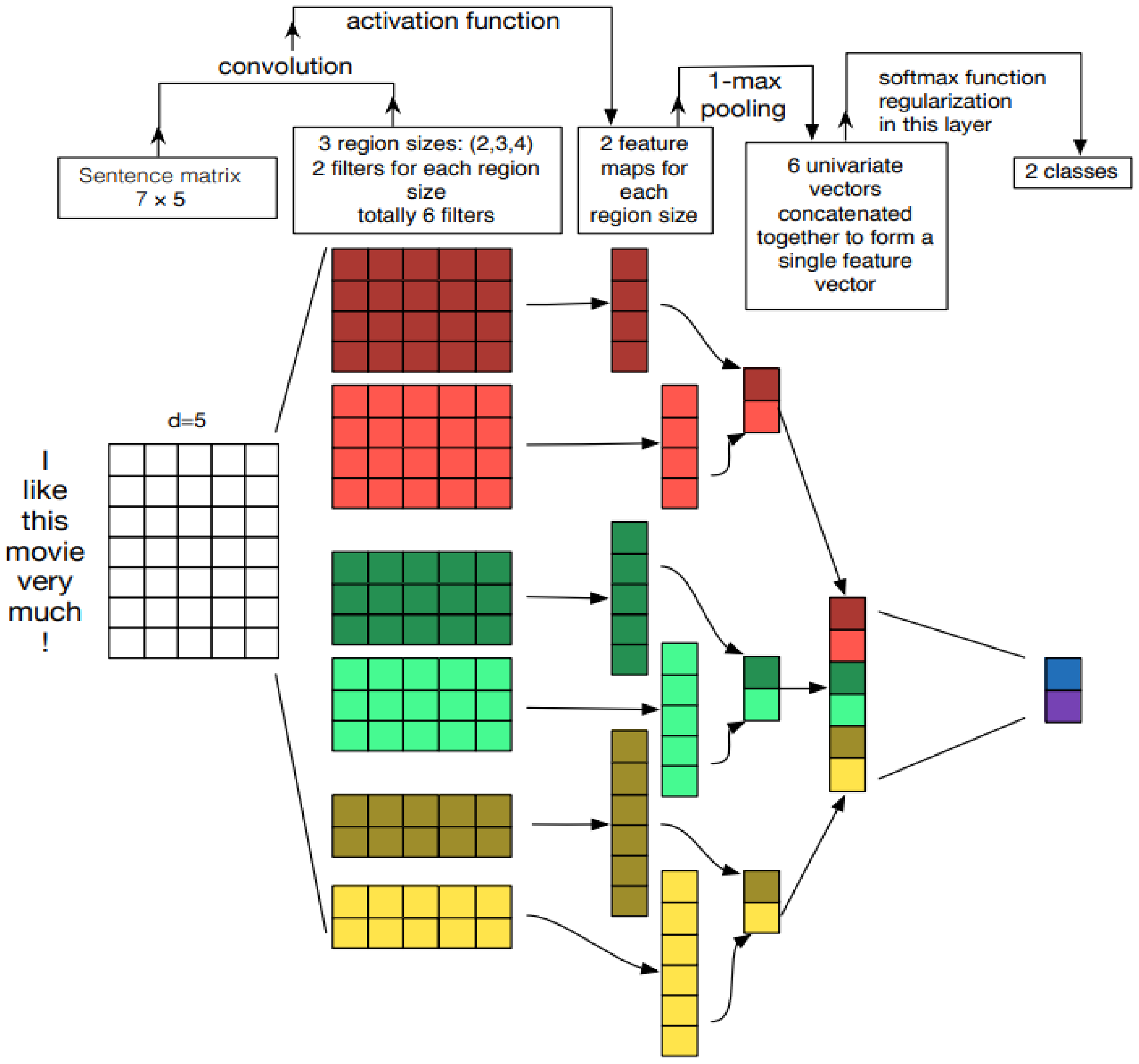
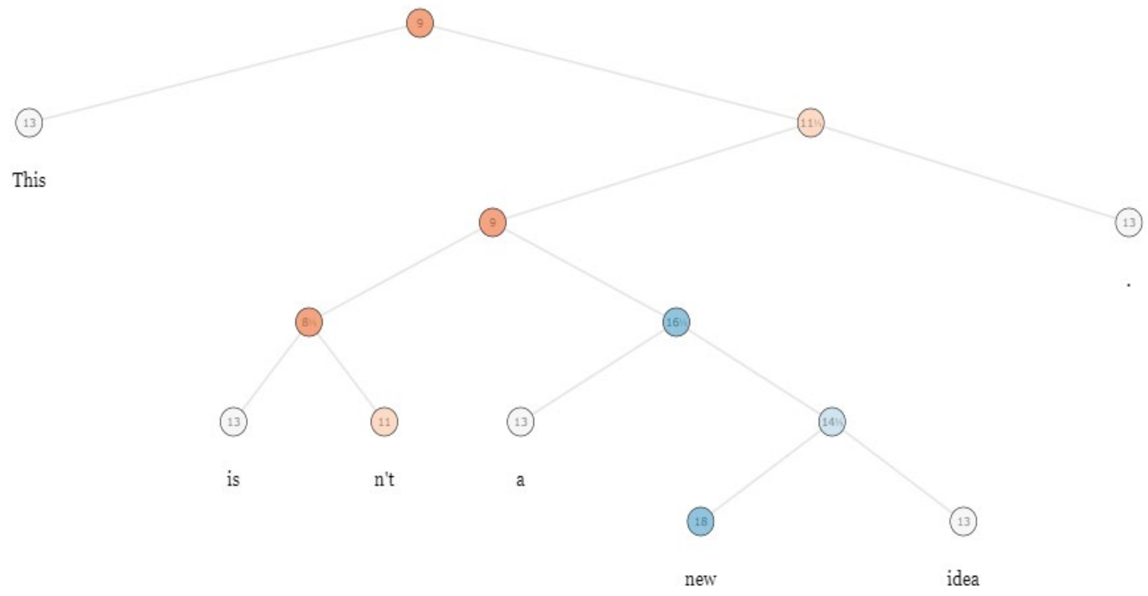


Illustration of a CNN architecture for sentence classification. We depict three filter region sizes: 2, 3 and 4, each of which has 2 filters. Filters perform convolutions on the sentence matrix and generate (variable-length) feature maps; 1-max pooling is performed over each map, i.e., the largest number from each feature map is recorded. Thus a univariate feature vector is generated from all six maps, and these 6 features are concatenated to form a feature vector for the penultimate layer. The final softmax layer then receives this feature vector as input and uses it to classify the sentence; here we assume binary classification and hence depict two possible output states.

From Zhang and Wallace, 2016.

Sentiment Analysis: Is a Flat Classification Good Enough?

- So far in sentiment analysis
  - Semantic representation of words.
  - Convolutional networks for sentiment classification.
- Stanford Sentiment Treebank: the first step towards "compositionality"



Explore more structures in https://nlp.stanford.edu/sentiment/

Each node has a sentiment score (1 to 25)

The project is much simpler: we try to predict the sentiment scores of "all" subtrees -> makes it a flat classification.

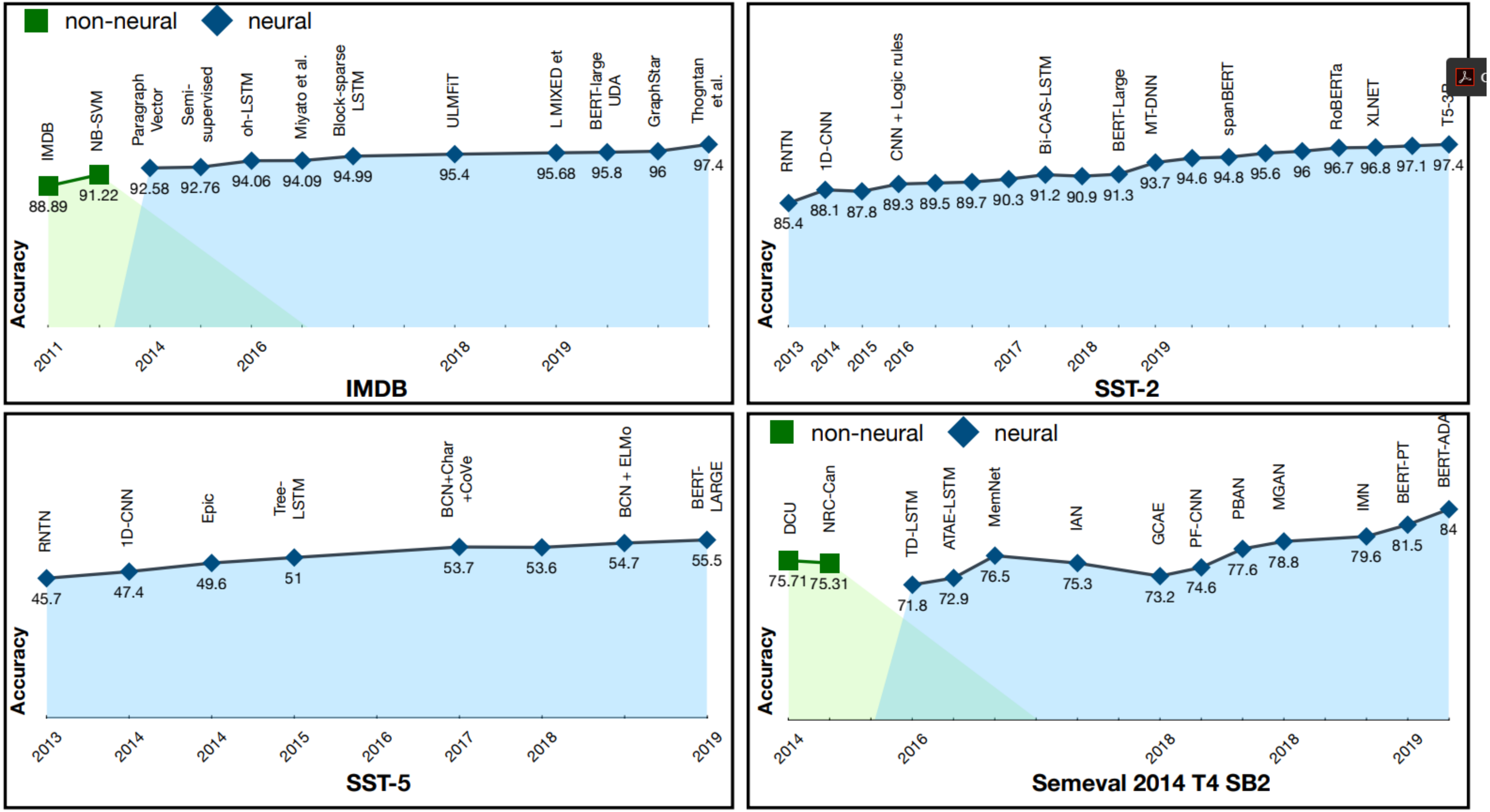
The actual goal is to predict the sentiment score of the final node, i.e., 9

This will require a tree structured network

Summary

- Supervised methods for sentiment classification.
- Very basic introduction to word convolutional network for sentiment classification (can be extended to any sentence classification task):
  - Representations of ``words" in a semantic space.
  - Building sentence representations from these words.
- Challenges for sentiment classification:
  - Sentence level: the challenge of compositionality.
  - Irony and sarcasm
  - Domain differences and context dependence

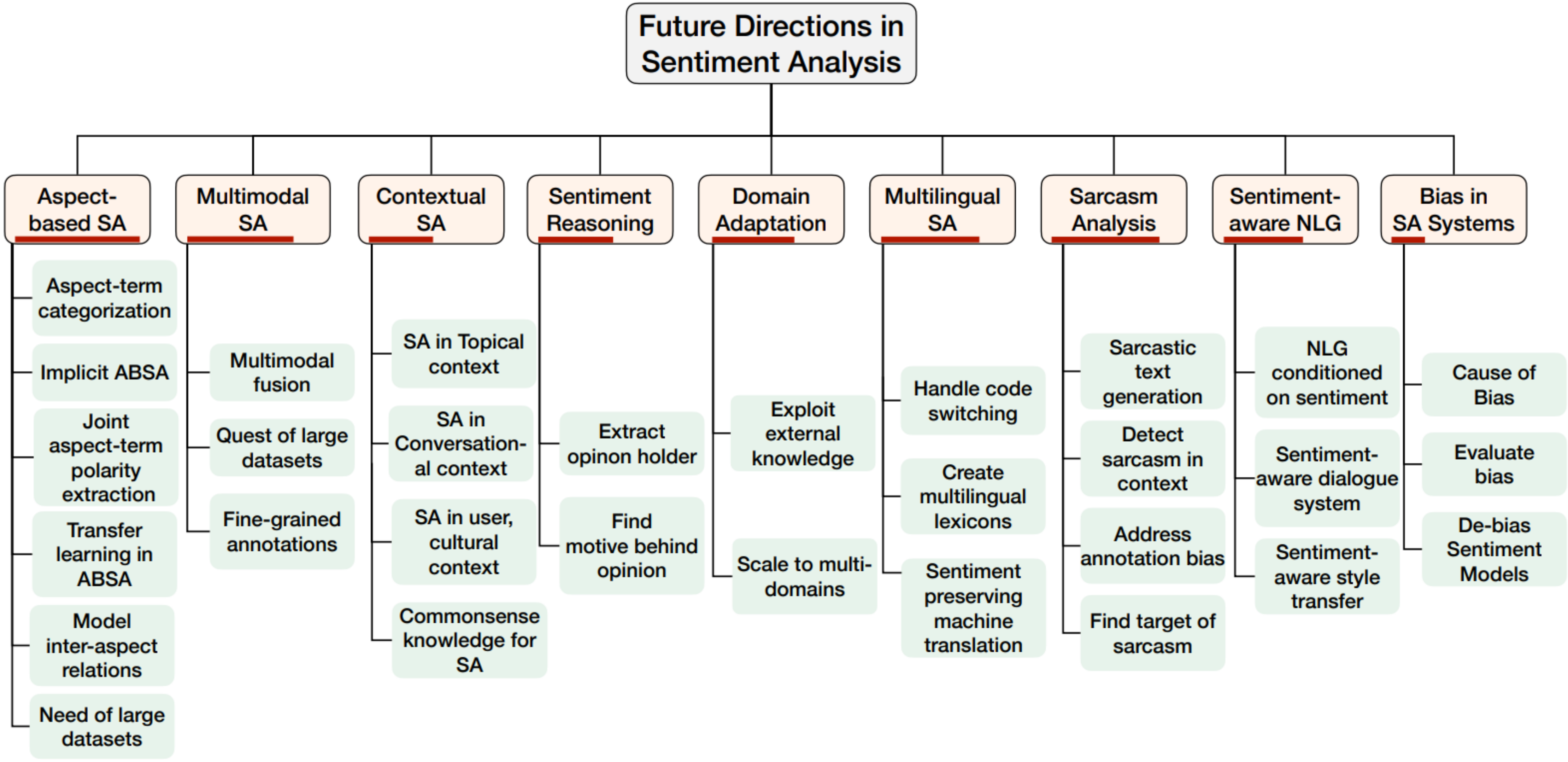
Sentiment Analysis: Open Research Areas



Current status of sentiment analysis datasets.

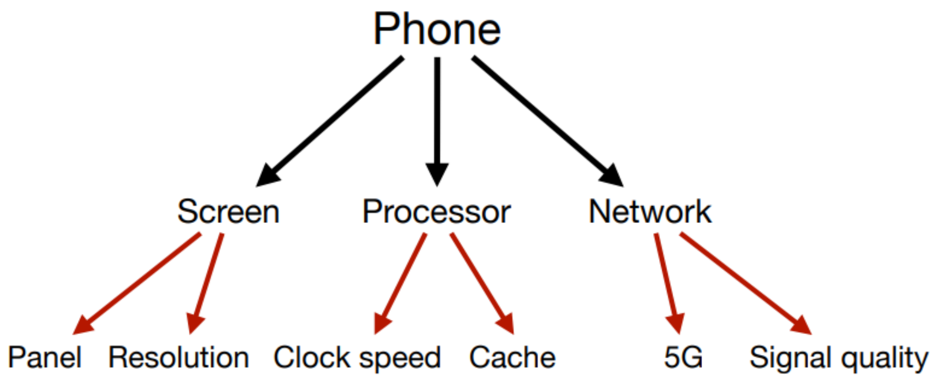
Poria et.al., 2021.

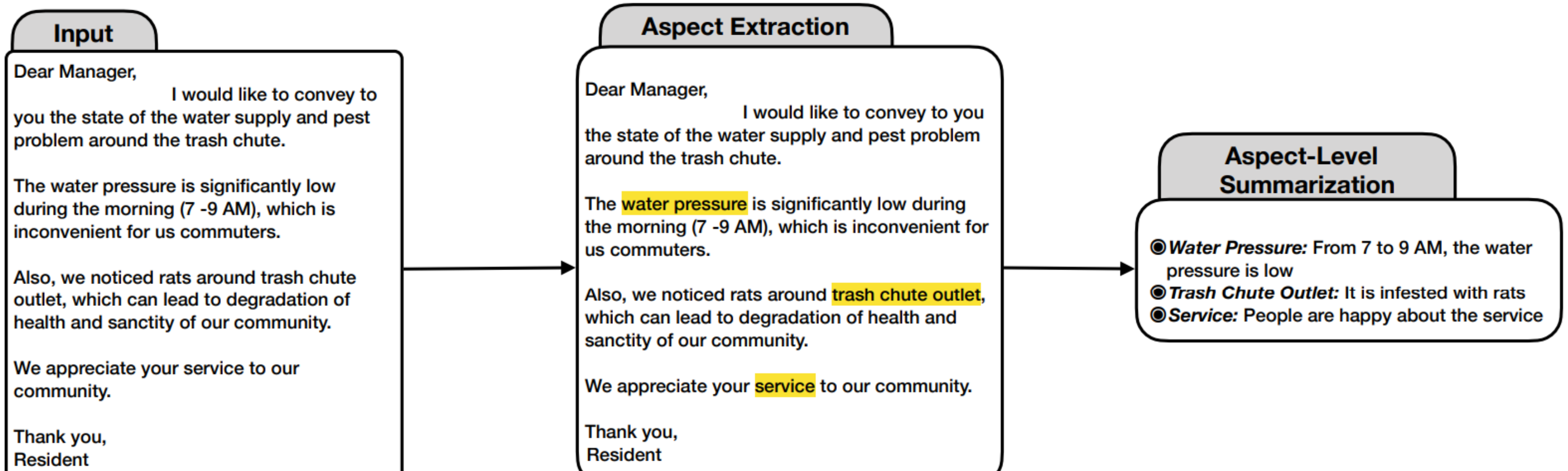
Sentiment Analysis: Future Directions



Aspect Based Sentiment Analysis

- "This actor is the only failure in an otherwise brilliant cast."
  - actor (negative opinion) and cast (positive opinion).
- Two sub-tasks:
  - Aspect extraction (similar to entity extraction) and Categorization/
  - Aspect level sentiment analysis.
  - With a proper aspect based sentiment analysis model, even summarization becomes Useful!





Multimodal Sentiment Analysis

Utterance: *"Become a drama critic!"*

Emotion: *Joy* **Sentiment:** *Positive*

Text	Audio	Visual
Ambiguous	Joyous tone	Smiling Face

Utterance: *"Great, now he is waving back"*

Emotion: *Disgust* **Sentiment:** *Negative*

Text	Audio	Visual
Positive/Joy	Flat tone	Frown

Utterance

1) **Chandler** :

Oh my god! You almost gave me a heart attack!

- **Text** : suggests fear or anger.
- **Audio** : animated tone
- **Video** : smirk, no sign of anxiety

2) **Sheldon** :

Its just a *privilege* to watch your mind at work.

- **Text** : suggests a compliment.
- **Audio** : neutral tone.
- **Video** : straight face.

Attend the QA session on Monday for clarifications and interesting discussions!