Text Classification and Sentiment analysis

NLP 2022

Rebecka Weegar, rebeckaw@dsv.su.se

February 2, 2022

What is text classification?

- Categorizing text
- Is this email a spam email or not?
- Who is the author of this text? Which language is it written in? Who is it about?
- Assign a label to a text where the label says something about the text content

What is text classification?

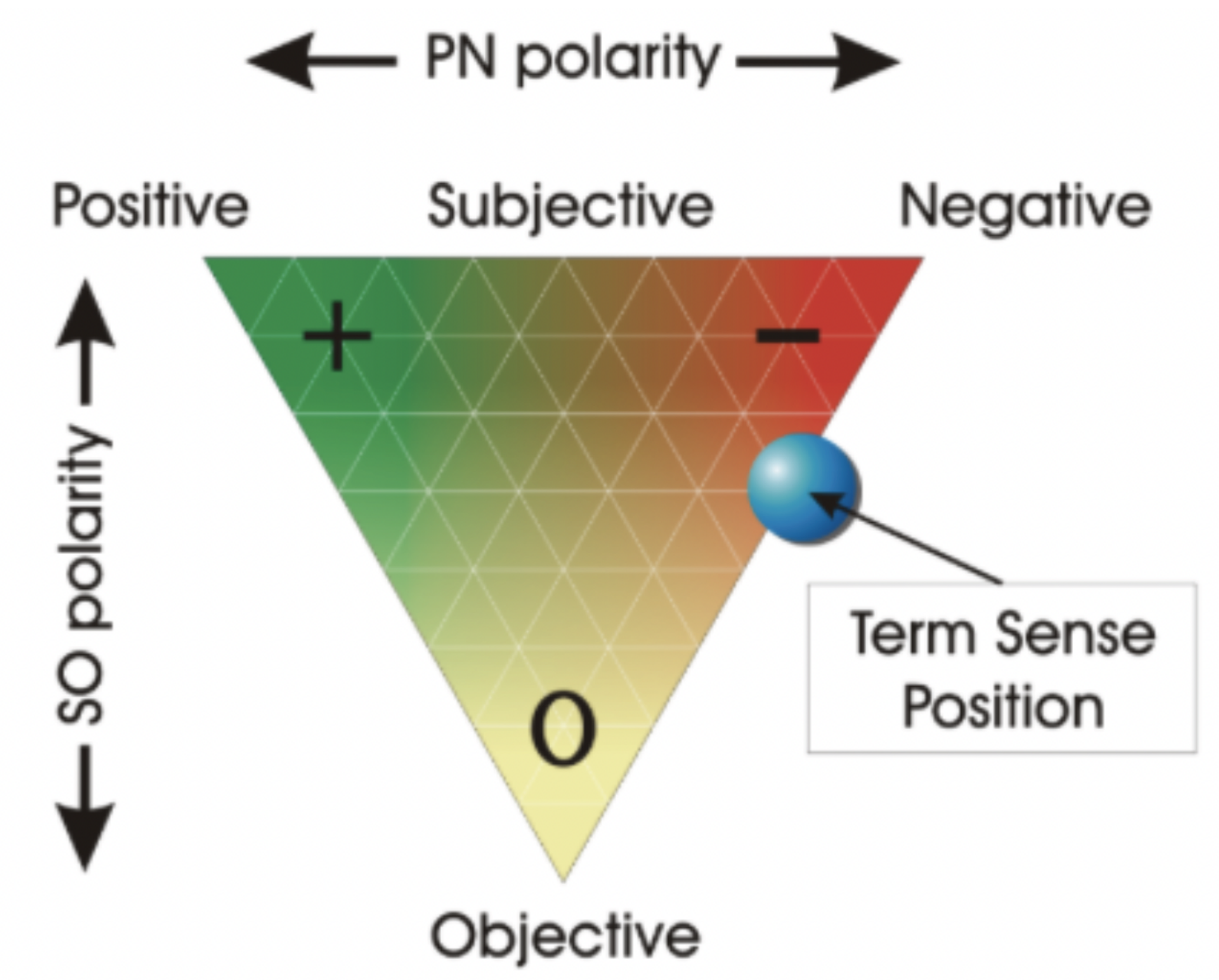
- Text classification is an important NLP task with diverse applications
- As supervised machine learning:
  - With a document X and a label y, the goal is to find the function f describing the relationship between the input and the output
- $y = f(X)$
- The label, y, depends on task and domain
- Clinical/Medical domain:
  - X: patient record, scientific article...
  - y: Diagnosis codes (ICD-10), MeSH, risk of adverse events

What is sentiment analysis?

- $y = f(X)$
- y says something about sentiment of the text (or emotion, attitude, opinion...)
- Commonly: $y \in \{POSITIVE, NEGATIVE\}$
- x: a word, a review, a tweet, a forum post, a book, a news article...
- Sentiment analysis is research field with many application areas:
  - What do people think about a specific product, movie, event?
  - Opinion mining: aggregate opinions about an opinion object
  - Assess general sentiment (and make predictions)
- We do not always agree on the correct classification y

Sentiment polarity

- The orientation of a statement, $\{POSITIVE, NEGATIVE\}$, is often called *polarity*
- $\{POSITIVE, NEGATIVE\}$ polarity is sometimes paired with a *subjectivity* dimension
- Subjective statements concerns the internal, attitudes, beliefs, state of mind
- Objective statements concerns the observable
- Can objective statements convey sentiment? (yes they can)



Esuli, A., and Fabrizio S. Sentiwordnet: A publicly available lexical resource for opinion mining. 2006.

Affective meaning in text

- Scherer typology of affective states:
  - Emotion, mood, interpersonal stance, attitude, personality traits
- Sentiment analysis focuses primarily on attitude, often as either negative or positive
- Part of understanding the meaning of a text is understanding the affective meaning
- Is this forum post written by a depressed person?
- Does this chat-bot interaction express frustration?
- How has attitude towards something changed over time?

Classifying text - Bag-of-words

- Bag-of-words is commonly used for text classification
- All words in a document, ignoring word order
- Preprocessing can be applied, such as removing stop words, lemmatization etc.
- Individual words can be represented by their counts, using weights such as tf-idf or as binary features

classifiers	1
introduce	1
will	8
help	3
analysing	1
effects	2
increase	1
interactions	3
matrix–vector	1
compositionality	3
considering	1
difference	4
standard	4
latter	1
tensor–based	1
higher	1
allows	1
smaller	1
constituents	2
the	757
a	381
artificial	9
intelligence	13

mitre	1
corner	1
take	3
cuffs	2
them	1
half	1
one	3
inside	3
sides	6
aligned	1
narrow	1
pull	5
understitch	2
baste	1
ends	1
remaining	2
notch	1
shoulder	1
press	13
either	1
pocket	2
the	42
a	9

Text classification with Naive Bayes

Naive Bayes is a probabilistic classifier: $\hat{\boldsymbol{c}} = \underset{\boldsymbol{c} \in \boldsymbol{C}}{\mathbf{argmax}} \boldsymbol{P}(\boldsymbol{c}|\boldsymbol{d})$

Select the class c of all classes C with the highest posterior probability given the document d

Bayes' rule for conditional probability:

$$Bayes' Rule : P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$

Bayes' rule for classification

Naive Bayes is a probabilistic classifier: $\hat{\boldsymbol{c}} = \underset{\boldsymbol{c} \in \boldsymbol{C}}{\mathbf{argmax}} \boldsymbol{P}(\boldsymbol{c}|\boldsymbol{d})$

Select the class c of all classes C with the highest posterior probability given the document d

Bayes' rule for conditional probability:

$$Bayes' Rule : P(c|d) = \frac{P(d|c)P(c)}{P(d)}$$

Bayes' rule for classification

$$Bayes' Rule : P(c|d) = \frac{\overbrace{P(d|c)}^{\text{LIKELIHOOD OF DOCUMENT GIVEN CLASS}}\overbrace{P(c)}^{\text{PRIOR CLASS PROBABILITY}}}{\underbrace{P(d)}_{\text{DOCUMENT PROBABILITY}}}$$

Training the Naive Bayes Classifier, Prior: P_c

Prior probabilities, probabilities of the different classes without seeing any text:

- N_{doc} : the number of documents in the training data
- N_c : the number of documents belonging to class c in the training data
- The (estimated) prior probability of class c: $\hat{P}_c = \frac{N_c}{N_{doc}}$

Training the Naive Bayes Classifier: $P(d|c)$

Is it possible to estimate $P(d|c)$ from training data?

- Conditional probability of a single word: $\hat{P}(w_i|c) = \frac{count(w_i,c)}{\sum_{w \in V} count(w,c)}$
- Concatenate all documents in one class to calculate individual word frequencies

But what about full documents?

- **Bag-of-words simplification**:
  $P(this\ book\ was\ really\ amazing) = P(really\ this\ book\ was\ amazing) = ...$
- **Naive simplification**: Assume all words in a document are independent:

$$P(this\ amazing\ book|POS) = P(this|POS) \cdot P(amazing|POS) \cdot P(book|POS)$$

Training the Naive Bayes Classifier: $P(d)$

$Bayes' Rule: P(c|d) = \frac{P(d|c)P(c)}{P(d)}$

- $P(d) = P(d|c) + P(d|\neg c)$
- Will be the same for all documents
- $\frac{P(d|c)P(c)}{P(d)} \propto P(d|c)P(c)$
- The denominator can be ignored in the calculations since we are interested in the class with the maximum probability (not the actual probability)

Training the Naive Bayes Classifier: $P(d|c)$

$$P(this\ amazing\ book|POS) = P(this|POS) \cdot P(amazing|POS) \cdot P(book|POS)$$

- Probability for unseen words?
- $count(\texttt{bad}|POS) = 0 \Rightarrow$ Probability of sentence $= 0$
- Laplace smoothing: Add 1 to all word counts

Training the Naive Bayes Classifier: $Log - space$

$$P(w_1, w_2, ..., w_m|class_j) = \prod_{i=1}^{m} P(w_i|class_j)$$

- Add logarithms of probabilities to avoid multiplying very small numbers:
- $a < b \Rightarrow log(a) < log(b)$
- $log(a \cdot b) = log(a) + log(b)$

$$P(w_1, w_2, ..., w_m|class_j) = \sum_{i=1}^{m} logP(w_i|class_j)$$

Naive Bayes Classifier

Putting it all together:

$$C_{NB} = \underset{c \in C}{\text{argmax}}\ log\ P(c) + \sum_i log\ P(w_i|c)$$

Select the class that maximized the conditional probability of the class given the document

- Instead of raw counts of word frequencies, binary counts can be more efficient as features for sentiment analysis
- It can be useful to handle negations:
  - The movie was boring VS The movie was *not* boring
- Using n-grams instead of words as features

Evaluation

- Common measures for evaluating text classification models: Precision, Recall, F1-score, Accuracy

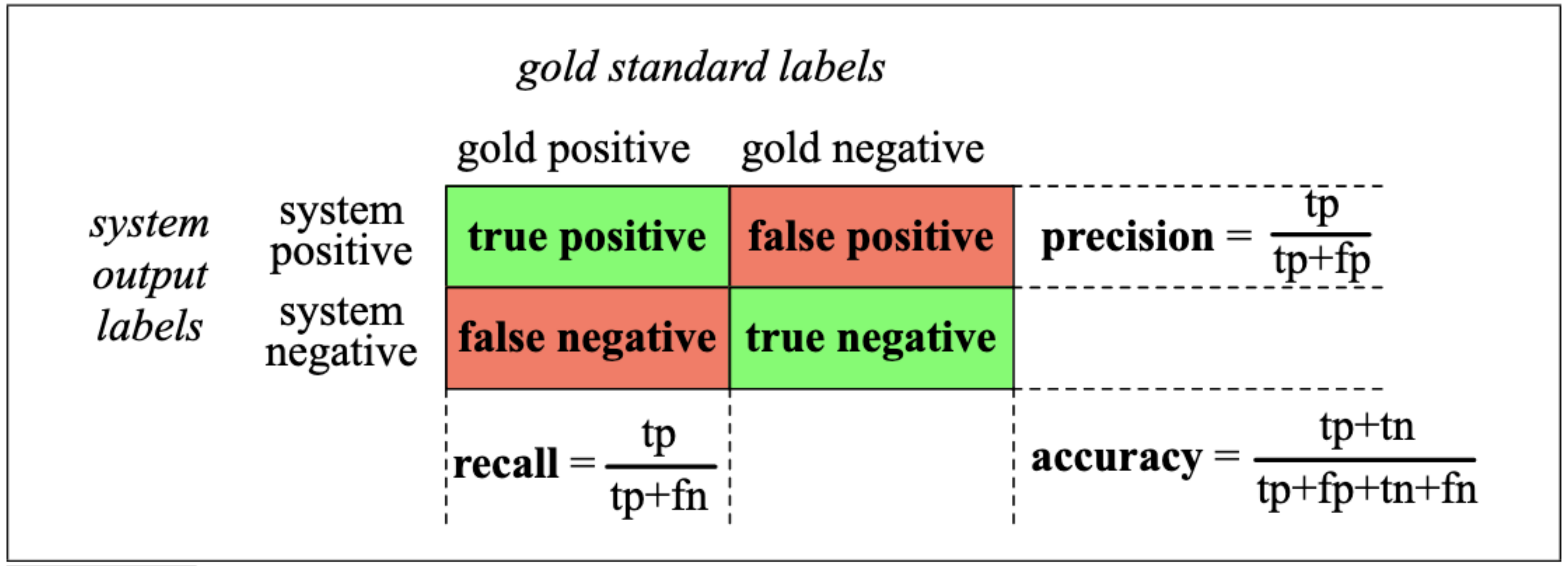


Figure 4.4 A confusion matrix for visualizing how well a binary classification system performs against gold standard labels.

Speech and Language Processing. Daniel Jurafsky & James H. Martin. (2021)

Evaluation - multi-class

- Micro-average: sum up all TP, FP, FN
- Macro-average: calculate per class and divide by number of classes

		gold labels urgent	normal	spam
system output	urgent	8	10	1
	normal	5	60	50
	spam	3	30	200

Class 1: Urgent

	true urgent	true not
system urgent	8	11
system not	8	340

$$\text{precision} = \frac{8}{8+11} = .42$$

Class 2: Normal

	true normal	true not
system normal	60	55
system not	40	212

$$\text{precision} = \frac{60}{60+55} = .52$$

Class 3: Spam

	true spam	true not
system spam	200	33
system not	51	83

$$\text{precision} = \frac{200}{200+33} = .86$$

Pooled

	true yes	true no
system yes	268	99
system no	99	635

$$\text{microaverage precision} = \frac{268}{268+99} = \mathbf{.73}$$

$$\text{macroaverage precision} = \frac{.42+.52+.86}{3} = \mathbf{.60}$$

Speech and Language Processing. Daniel Jurafsky & James H. Martin. (2021) **There will be a separate lecture on evaluation of all things NLP**

Sentiment analysis - levels

- How long is a text?
- A news article, a book, a tweet, a word?
- Challenges with short text: very little information available
- Challenges with longer texts: can express several different sentiments

Filmrecension: Banal intrig men estetisk fest i nya adaptionen av rymdeposet Dune

UPPDATERAD 15 SEPTEMBER 2021 PUBLICERAD 15 SEPTEMBER 2021

Om intrigen ter sig sliten och banal så är det estetiska desto mer imponerande och bör absolut åtnjutas i biograf. Det tycker Fredrik Sahlin efter att ha sett Denis Villeneuves omtalade och coronaförsenade adaption av Frank Herberts påstått ofilmbara epos.

Frank Herberts epos har på grund av dess komplexa berättelse sagts vara ofilmbar, men det kan i och för sig ha varit ett rykte som David Lynch spred ut efter det att han kraschlandat med sin kalkonversion från 1984 – en film som han själv hatar så mycket att han fortfarande vägrar prata om den. Även den gamle chilenska sci-fi-mästaren Alejandro Jodorowsky gjorde ett tappert försök till adaption i ett verk som, om det hade blivit av, hade klockat in på 15 timmar.

Nå, för den som inte har läst förlagan, och bara kan döma utefter filmerna, ter sig Dunes story inte så fasligt komplicerad. I alla fall inte grundförutsättningarna:

Återigen ser vi den utvalde unge mannen som först inte vill leva upp till sitt ansvar men sedan faller till föga, härdas och går i sin faders och profetiornas spår. Precis som Jesus, Kung Arthur, Luke Skywalker och alla de andra Kristus-figurerna.

Det växer en hel del spindelväv på den premissen...

Vi har också en maktfullkomlig kejsare som leder det hårt härskande *Imperiet*, en ond släkt kallad Harkonnen och deras motståndare, de ärbara Atreides – och jodå, även i den här galaxen ”far, far away” finns det en psykokinetisk *Kraft* som bara ett fåtal behärskar.

Det gäller dock att påminna sig själv – och ungdomar i ens omgivning som ropar ”ripp-off!” – att förlagan faktiskt kom ett drygt decennium före det att George Lucas började dra in astronomiska summor på sina rymdsagor.

Om nu intrigen ter sig sliten och banal (förutom möjligen en anti-kolonialistisk ådra som dock slår knut på sig själv) så är det visuella desto mer imponerande och bör absolut åtnjutas i biograf. Och sedd i en IMAX-salong ger Dune en ännu maffigare, helomslutande matinékänsla. Varje dollar finns, som man brukar säga, på duken.

Det är oerhört snyggt, mäktigt och alla andra ord man kan tänka sig för att (fruktlöst) försöka förmedla en estetisk fest som denna. Hans Zimmers som vanligt imposanta musik, den fantasifulla produktionsdesignen och det suggestiva fotot – allt sitter där det ska och förmedlar storhet och ödesmättad tyngd; en illusion av att det vi tittar på är av vikt.

”En Star Wars för vuxna” lär den regisserande kanadicken Denis Villeneuve (Arrival, Blade runner 2049) själv ha sagt om sitt verk, och jo, vi slipper alla infantila muppar och skojfriska inslag (som den evigt irriterande roboten C-3PO). Här är det istället gravallvar och melankoli som gäller, och högstämda tal om mandom, mod och morske män – vilket i och för sig kan bli ofrivilligt skrattretande.

Mest rörande är Stellan Skarsgård som med köttig ”fat suit” gör ärkeskurken Baron Vladimir Harkonnen till en slemmig och lite ömklig heffaklump – han tillåter sig till och med att slänga in några referenser till Marlon Brandos plågade patriark Överste Kurtz i Apocalypse Now.

Denis Villeneuve bestämde tidigt att inte, som David Lynch, klämma in hela berättelsen i en film, eftersom han inte ville våldföra sig för mycket på Herberts verk. Vilket i och för sig är klokt. Å andra sidan känns det som ett rejält antiklimax när eftertexterna börjar rulla mitt i det äventyr som nu går in en halvtidsvila som kommer att vara i åratal.

https://www.svt.se/kultur/film/filmrecension-dune-svt-kulturnyheterna-fredrik-sahlin

Sentiment analysis - aspect level

- There can be several aspects of the same opinion object
- Different sentiments can be expressed for different aspects in the same text
- Find (and group) aspects and their polarity

```
The plot is worn out, banal
The story ... not too complicated
The premise is covered in cobwebs
melancholia ... unintentionally laughable
the aesthetic is all the more impressive
The visuals should be enjoyed at a movie theatre
Every dollar is on the screen
```

Filmrecension: Banal intrig men estetisk fest i nya adaptionen av rymdeposet Dune

UPPDATERAD 15 SEPTEMBER 2021 PUBLICERAD 15 SEPTEMBER 2021

Om intrigen ter sig sliten och banal så är det estetiska desto mer imponerande och bör absolut åtnjutas i biograf. Det tycker Fredrik Sahlin efter att ha sett Denis Villeneuves omtalade och coronaförsenade adaption av Frank Herberts påstått ofilmbara epos.

Frank Herberts epos har på grund av dess komplexa berättelse sagts vara ofilmbar, men det kan i och för sig ha varit ett rykte som David Lynch spred ut efter det att han kraschlandat med sin kalkonversion från 1984 – en film som han själv hatar så mycket att han fortfarande vägrar prata om den. Även den gamle chilenska sci-fi-mästaren Alejandro Jodorowsky gjorde ett tappert försök till adaption i ett verk som, om det hade blivit av, hade klockat in på 15 timmar.

Nå, för den som inte har läst förlagan, och bara kan döma utefter filmerna, ter sig Dunes story inte så fasligt komplicerad. I alla fall inte grundförutsättningarna:

Återigen ser vi den utvalde unge mannen som först inte vill leva upp till sitt ansvar men sedan faller till föga, härdas och går i sin faders och profetiornas spår. Precis som Jesus, Kung Arthur, Luke Skywalker och alla de andra Kristus-figurerna.

Det växer en hel del spindelväv på den premissen...

Vi har också en maktfullkomlig kejsare som leder det hårt härskande *Imperiet*, en ond släkt kallad Harkonnen och deras motståndare, de ärbara Atreides – och jodå, även i den här galaxen "far, far away" finns det en psykokinetisk *Kraft* som bara ett fåtal behärskar.

Det gäller dock att påminna sig själv – och ungdomar i ens omgivning som ropar "ripp-off!" – att förlagan faktiskt kom ett drygt decennium före det att George Lucas började dra in astronomiska summor på sina rymdsagor.

Om nu intrigen ter sig sliten och banal (förutom möjligen en anti-kolonialistisk ådra som dock slår knut på sig själv) så är det visuella desto mer imponerande och bör absolut åtnjutas i biograf. Och sedd i en IMAX-salong ger Dune en ännu maffigare, helomslutande matinékänsla. Varje dollar finns, som man brukar säga, på duken.

Det är oerhört snyggt, mäktigt och alla andra ord man kan tänka sig för att (fruktlöst) försöka förmedla en estetisk fest som denna. Hans Zimmers som vanligt imposanta musik, den fantasifulla produktionsdesignen och det suggestiva fotot – allt sitter där det ska och förmedlar storhet och ödesmättad tyngd; en illusion av att det vi tittar på är av vikt.

"En Star Wars för vuxna" lär den regisserande kanadicken Denis Villeneuve (Arrival, Blade runner 2049) själv ha sagt om sitt verk, och jo, vi slipper alla infantila muppar och skojfriska inslag (som den evigt irriterande roboten C-3PO). Här är det istället gravallvar och melankoli som gäller, och högstämda tal om mandom, mod och morske män – vilket i och för sig kan bli ofrivilligt skrattretande.

Mest rörande är Stellan Skarsgård som med köttig "fat suit" gör ärkeskurken Baron Vladimir Harkonnen till en slemmig och lite ömklig heffaklump – han tillåter sig till och med att slänga in några referenser till Marlon Brandos plågade patriark Överste Kurtz i Apocalypse Now.

Denis Villeneuve bestämde tidigt att inte, som David Lynch, klämma in hela berättelsen i en film, eftersom han inte ville våldföra sig för mycket på Herberts verk. Vilket i och för sig är klokt. Å andra sidan känns det som ett rejält antiklimax när eftertexterna börjar rulla mitt i det äventyr som nu går in en halvtidsvila som kommer att vara i åratal.

https://www.svt.se/kultur/film/filmrecension-dune-svt-kulturnyheterna-fredrik-sahlin

Sentiment analysis - aspect level

`The plot is worn out, banal`
`The story ... not too complicated`
`The premise is covered in cobwebs`
`melancholia ... unintentionally laughable`
`the aesthetic is all the more impressive`
`The visuals should be enjoyed at a movie theatre`
`Every dollar is on the screen`

Explicit aspects:

{Plot, Story, Premise}
{Visuals, Aesthetic}

Some aspects are implicit

Domain dependent: `predictable`

Sentiment analysis - aspect level

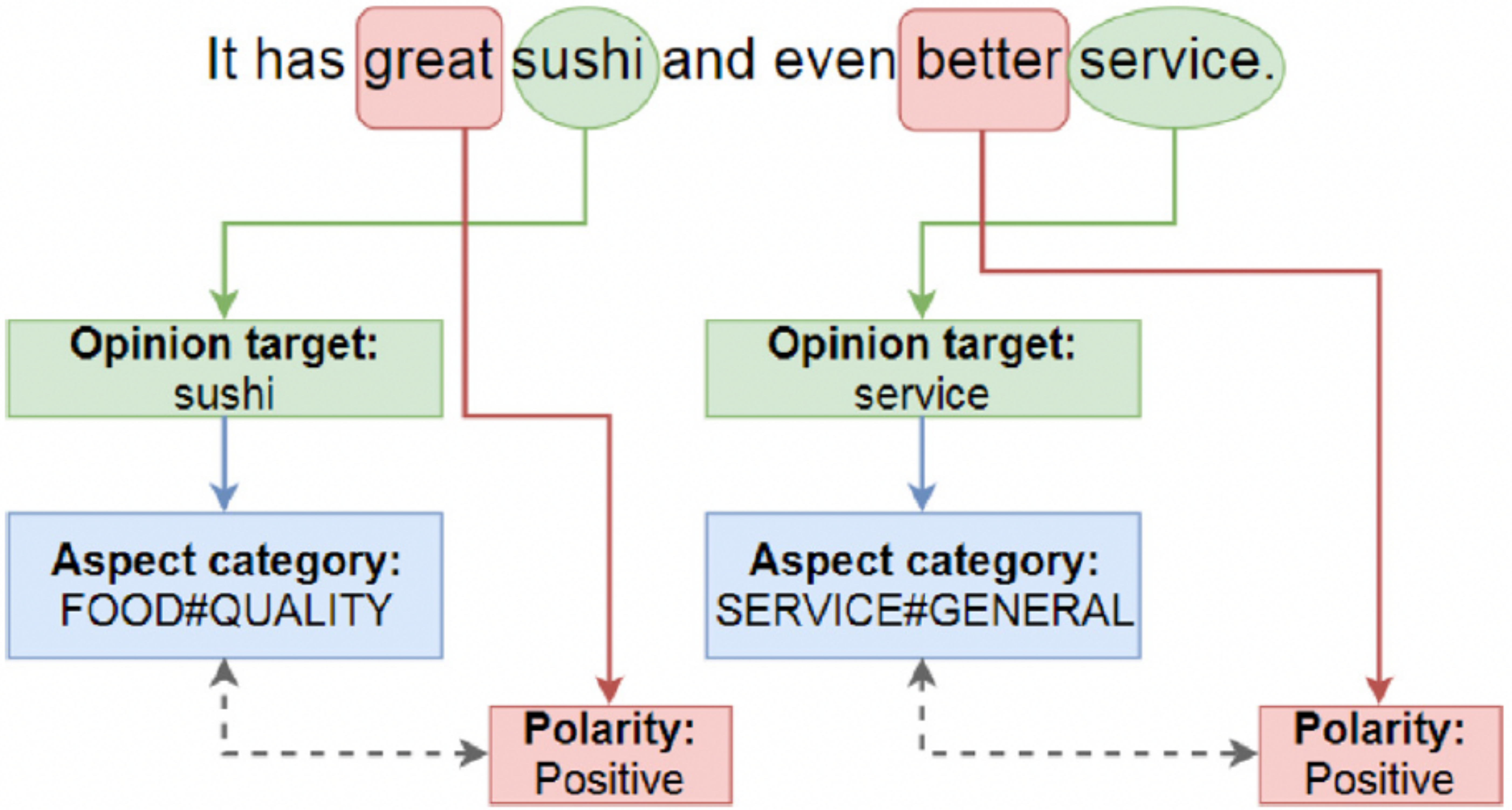


Fig. 1. Three tasks of ABSA in a sample sentence from SemEval ABSA dataset 2016. The sentence has two opinion targets: sushi & service. The category of "sushi" is "Food", with the attribute being "Quality" and polarity "Positive". The category is "Service", with an attribute of "General" and polarity of "Positive".

Do et al. Deep learning for aspect-based sentiment analysis: a comparative review. (2019)

Sentiment bearing words

- Individual words can express sentiment
- Valance: how pleasant something is (postive-negative) scale
- Arousal: the intensity of an emotion
- Adjectives and adverbs: beautiful, terribly, strong
- Nouns and verbs: hate, love, disaster
- Knowing about the sentiment of individual words can help sentiment classification

Sentiment lexicons

- Lexicons specifically created for describing sentiment values of words
- SentiWordNet `https://github.com/aesuli/SentiWordNet` POS-NEG, OBJECTIVE-SUBJECTIVE
- Binary: Words are either labelled as positive or negative
- Ordered:
  - Highest to lowest valance (pleasantness)
  - By emotional intensity

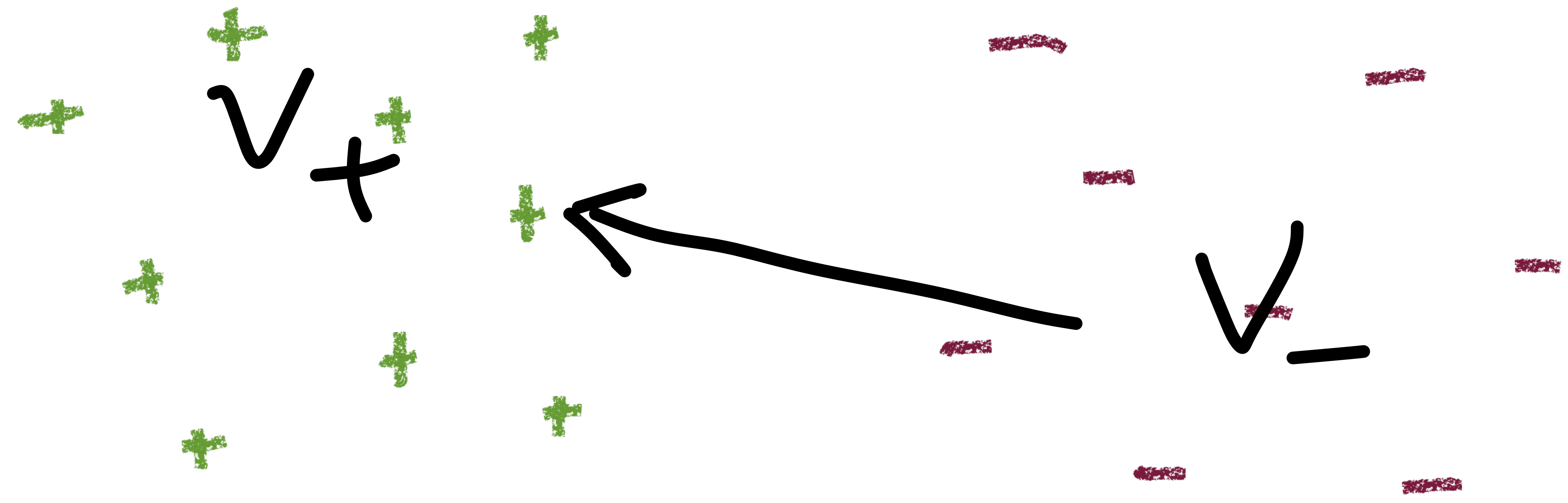
Word	Anger	Word	Fear	Word	Joy	Word	Sadness
outraged	0.964	*horror*	0.923	*sohappy*	0.868	*sad*	0.844
brutality	0.959	*horrified*	0.922	*superb*	0.864	*suffering*	0.844
satanic	0.828	*hellish*	0.828	*cheered*	0.773	*guilt*	0.750

Word Affect Intensities, Saif M. Mohammad (2018)

Creating sentiment lexicons - manual labelling

- Manual labelling of words can be useful for a language or domain without any existing lexical resources for sentiment analysis
- Sentiments categories are often subjective, even for single words
- How positive is *Happy, Boring, Summer, Olympic*?
- Annotations by a group of annotators more reliable
- Asking annotators to compare and rank words
- Which labels are most useful? General or specific labels?

Creating sentiment lexicons - sentiment axis



- A semi-supervised method
- Start with a small set of seed words
- The sentiment of each seed word is known
- Find the centroid of the embeddings of the negative and positive seed words
- Sentiment axis $V_{pos} - V_{neg}$
- Extend lexicon with words according to their cosine similarity to the sentiment axis

Creating sentiment lexicons from categorized texts

- Infer sentiment values of words based on reviews with scores
- Positive words occur in reviews with high scores
- Compare the probability of a word occurring for each score to determine the sentiment values of different words
- Similar methods can be used to find words related to other categories of text:
- Which words do leaders of different political parties use?

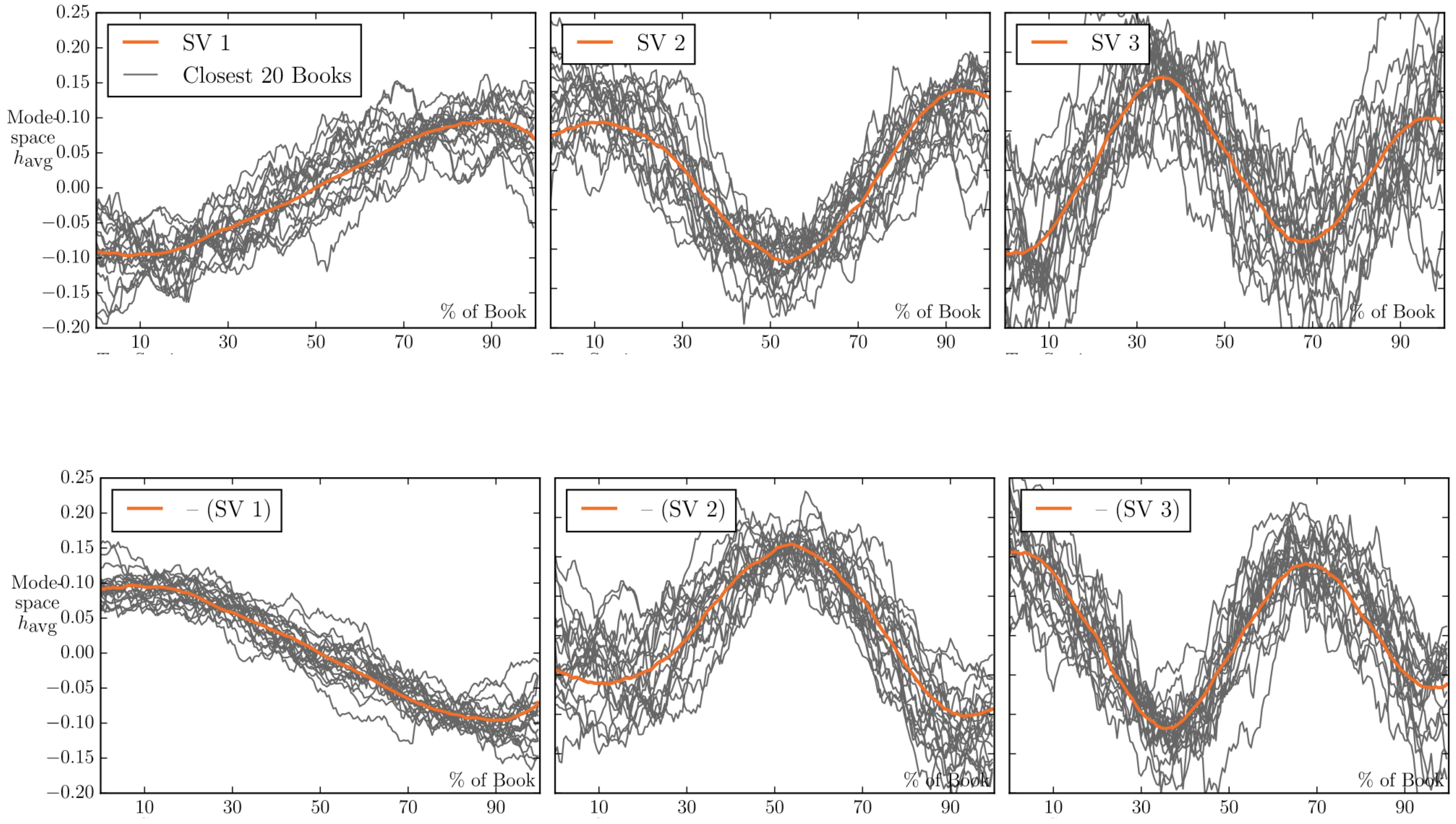
Using sentiment lexicons

- Most useful when only small training sets are available
- New languages or new domains might not have training data available
- Compare the proportions of positive and negative words in a text to classify its sentiment...
- or use the proportions or counts as features for a supervised classifier
- or use the sentiment values as weights for words in a bag-of-words model

Sentiment analysis applied to literature

- Finding the emotional arcs of stories
- Large sliding window applied to fiction to determine the emotion at different points

- Rags to riches' (rise)
- Tragedy (fall).
- Man in a hole (fall-rise)
- Icarus' (rise-fall)
- Cinderella (rise-fall-rise)
- Oedipus (fall-rise-fall).

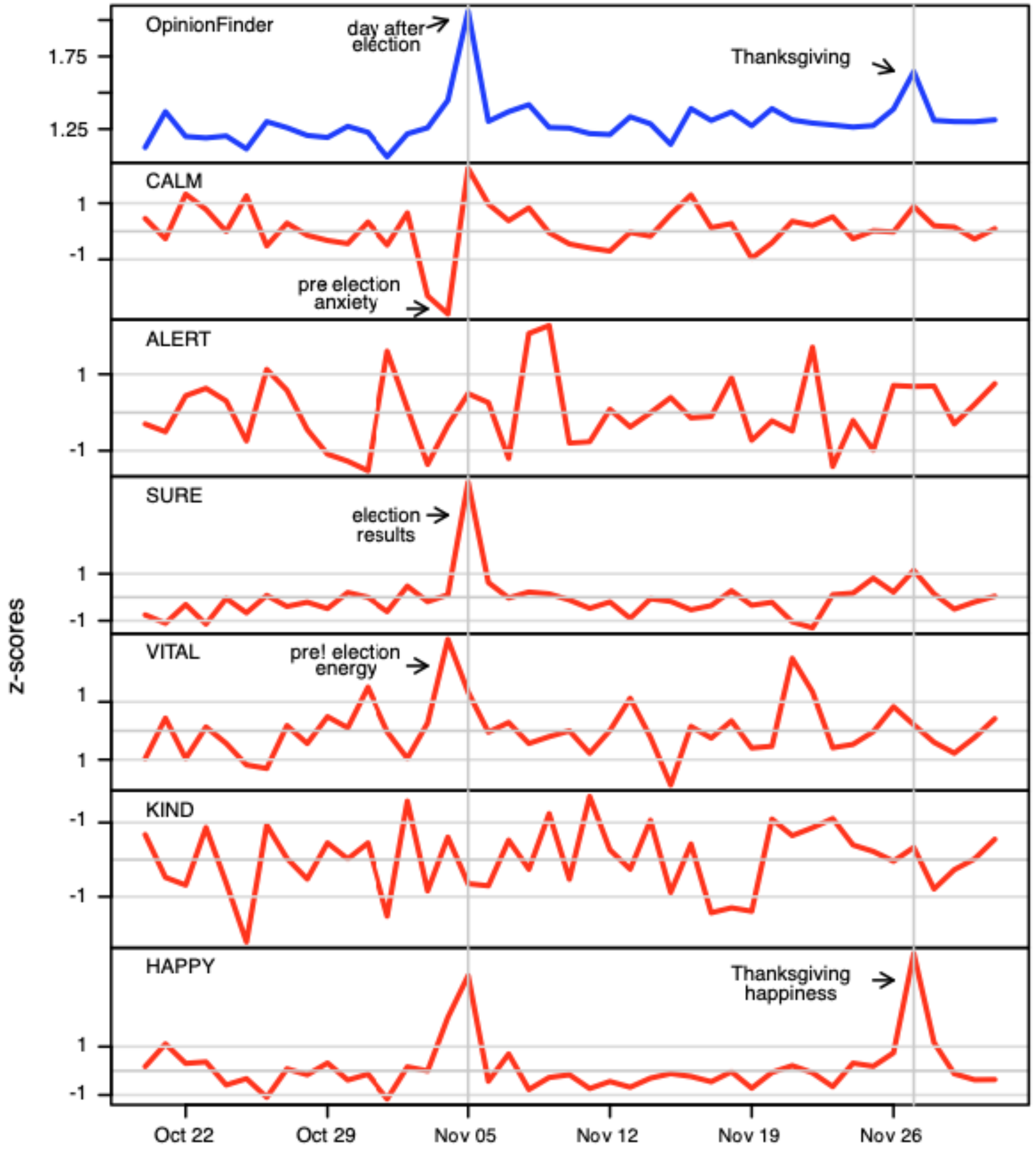


The emotional arcs of stories are dominated by six basic shapes. Reagan et al., (2016)

Measuring general mood

- What do people in general think or feel?
- Analyzing large amounts of twitter text
- Positive/Negative
- + 6 different moods, Calm, Alert, Sure, Vital, Kind, Happy
- Adding mood as a feature improved stock market predictions

Bollen et al. Twitter mood predicts the stock market (2011)

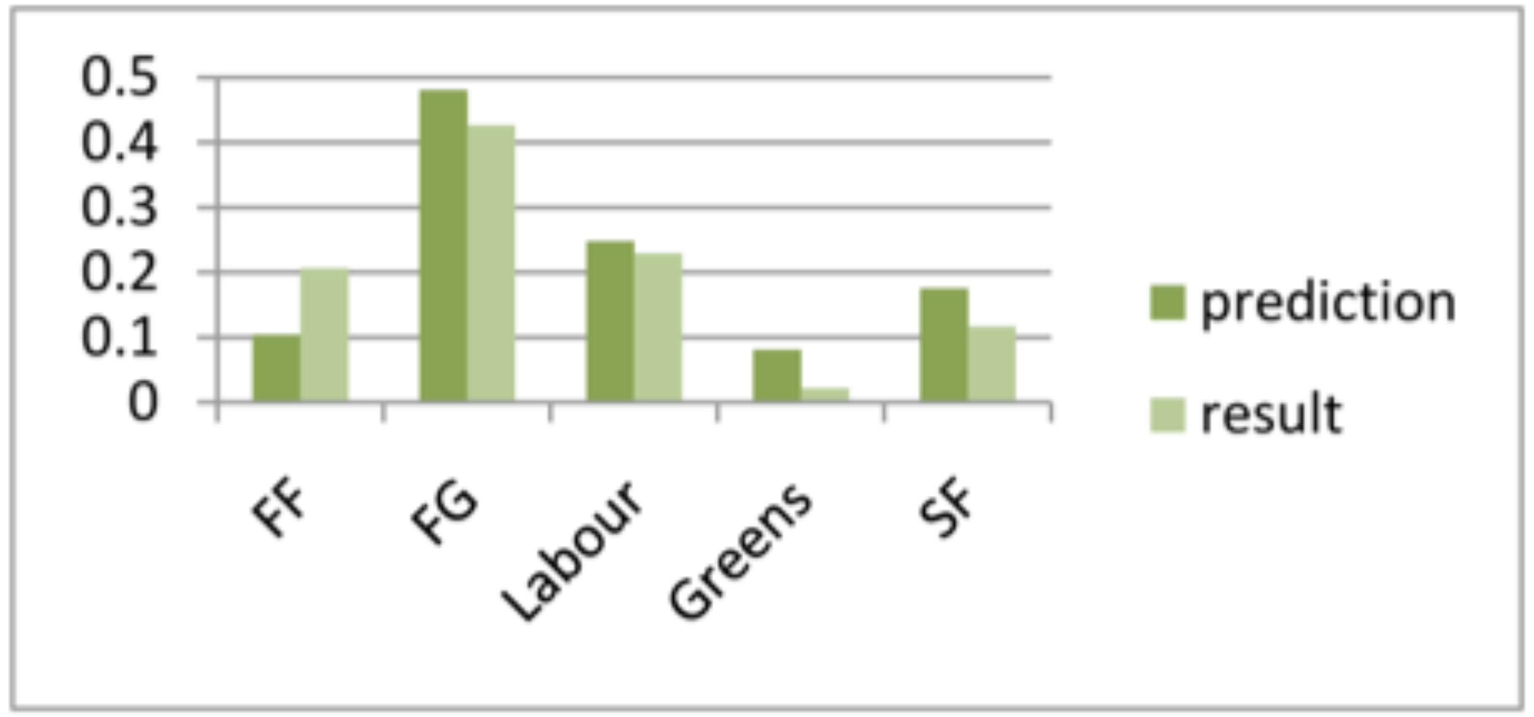


Hedonometer

Measuring average happiness for twitter: hedonometer.org

Prediction election outcomes

- Using social media to monitor/measure political sentiment
- Using this information for election outcome prediction?
- Tweets mentioning the election or political parties
- the volume of tweets + the sentiment of the tweets
- Positive, negative, neutral
- Challenges: noisy data, mostly negative sentiment, representativeness



On using Twitter to monitor political sentiment and predict election results. Bermingham, Adam and Smeaton, Alan

Sentiment analysis - negative sentiments

- Sentiment analysis can be used for monitoring language use
- Monitoring comment fields or forums to filter comments
- Detecting opinion spam, "troll detection", propaganda
- For opinion mining, are the opinions genuine?

Assigning clinical codes to text - example

- ICD-10 codes
- Example: $\sim$ 6000 gastrointestinal discharge summaries
- 263 Different codes

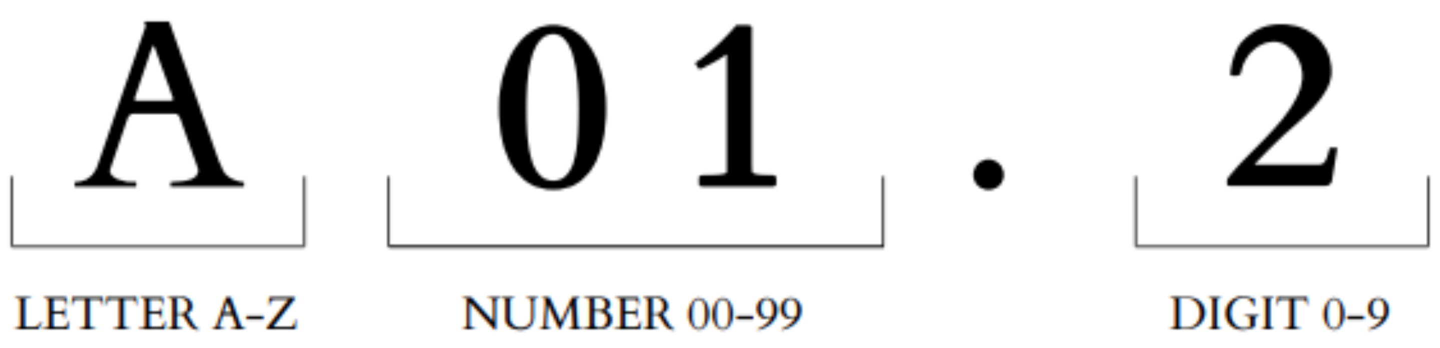


Figure 1: The anatomy of ICD codes.

	Macro			*Micro*		
Classifier	*P*	*R*	F_1	*P*	*R*	F_1
KB-BERT	0.67	0.55	0.60	0.87	0.77	0.82
SVM	0.76	0.33	0.41	0.90	0.61	0.72
DT	0.54	0.50	0.52	0.72	0.69	0.71
KNN	0.63	0.41	0.48	0.79	0.64	0.71

Table 6: Combined scores for the **Blocks** data set during the 10-fold cross-validation.

	Macro			*Micro*		
Classifier	*P*	*R*	F_1	*P*	*R*	F_1
KB-BERT	0.00	0.00	0.00	0.00	0.00	0.00
SVM	0.06	0.01	0.01	0.85	0.05	0.10
DT	0.10	0.09	0.09	0.30	0.28	0.29
KNN	0.11	0.03	0.05	0.55	0.17	0.26

Table 5: Combined scores for the **Full codes** data set during the 10-fold cross-validation.

Multi-label Diagnosis Classification of Swedish Discharge Summaries–ICD-10 Code Assignment Using KB-BERT, Remmer S., Lamproudis A., and Dalianis H.

Text classification in education

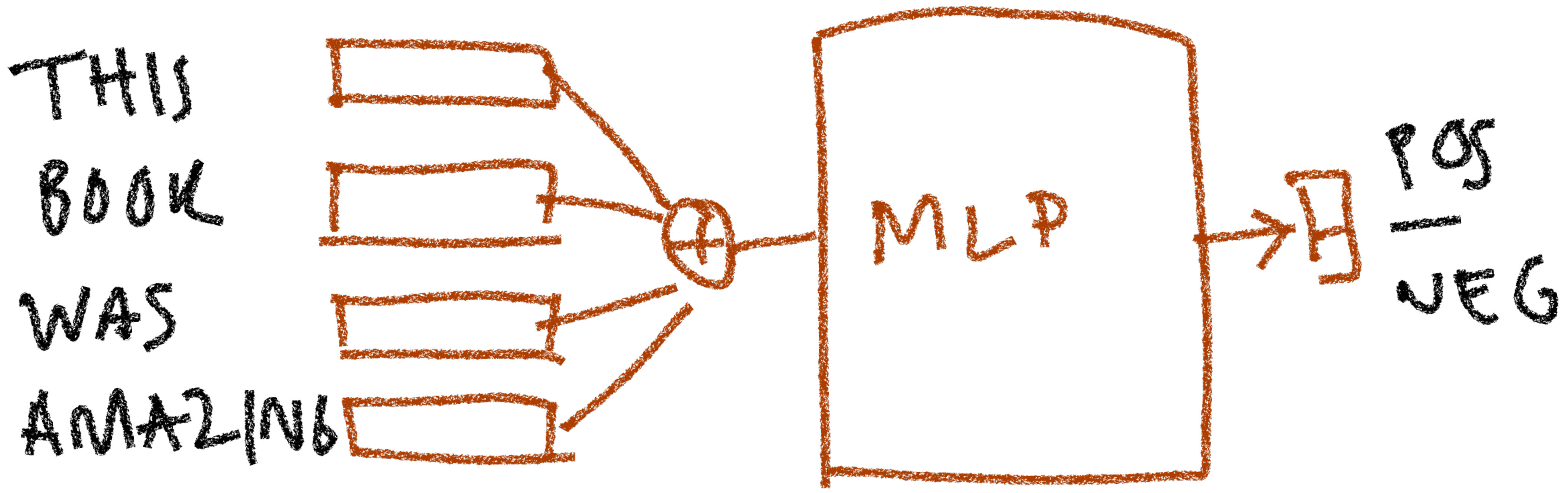
- Automated grading or assessment of student texts
- Short answer scoring
  - Similarity-based: how similar is this answer to a model answer
  - Supervised learning: train a classifier on a labelled set of answers
  - Label propagation: find clusters similar answer and grade only one answer for each cluster
  - Main focus is content
- Essay scoring
  - Structure and style are also important
  - Argumentative structure
  - Text quality measures are useful as features

Methods for text classification

- There have been several (overlapping) method paradigms for text classification
- Rules and dictionaries, conventional/shallow machine learning, "general" deep learning architectures, transformers/language models
- Most of the current research on sentiment analysis and other types of text classification applies deep learning of some kind
- Drawbacks: limited interpretability, most useful for very large data sets, large carbon footprint
- Several deep learning architectures can be useful for classifying texts
- A simplified model:
  - Input layer, Hidden layers, Output layer

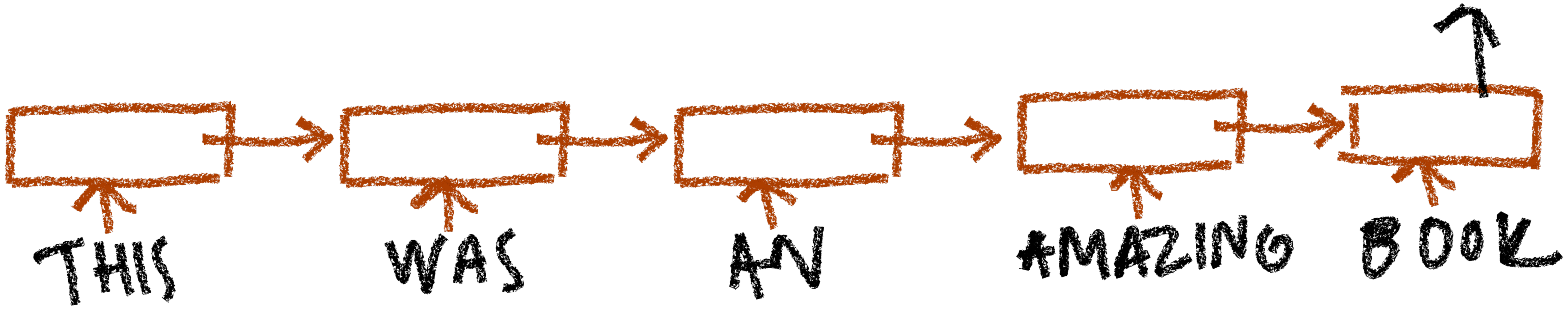
Deep learning architectures for text classification and sentiment analysis

- Feed forward networks
- **Text as a Bag-of-words**: Input layer values can be summed or averaged
- Input layer: Word embedding
- Hidden layers: Multilevel perceptron
- Output layer: shallow ML or softmax



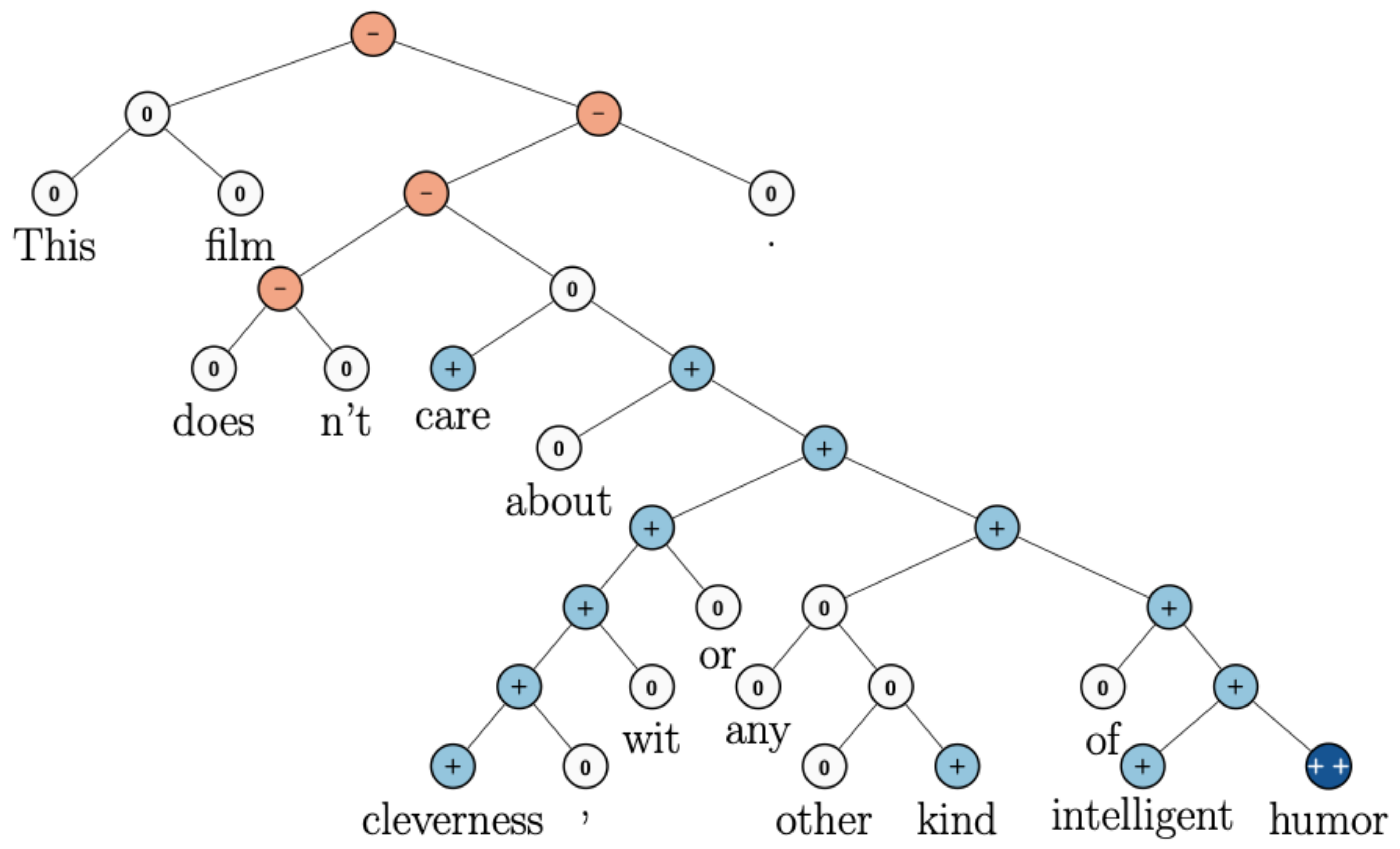
Deep learning architectures for text classification and sentiment analysis

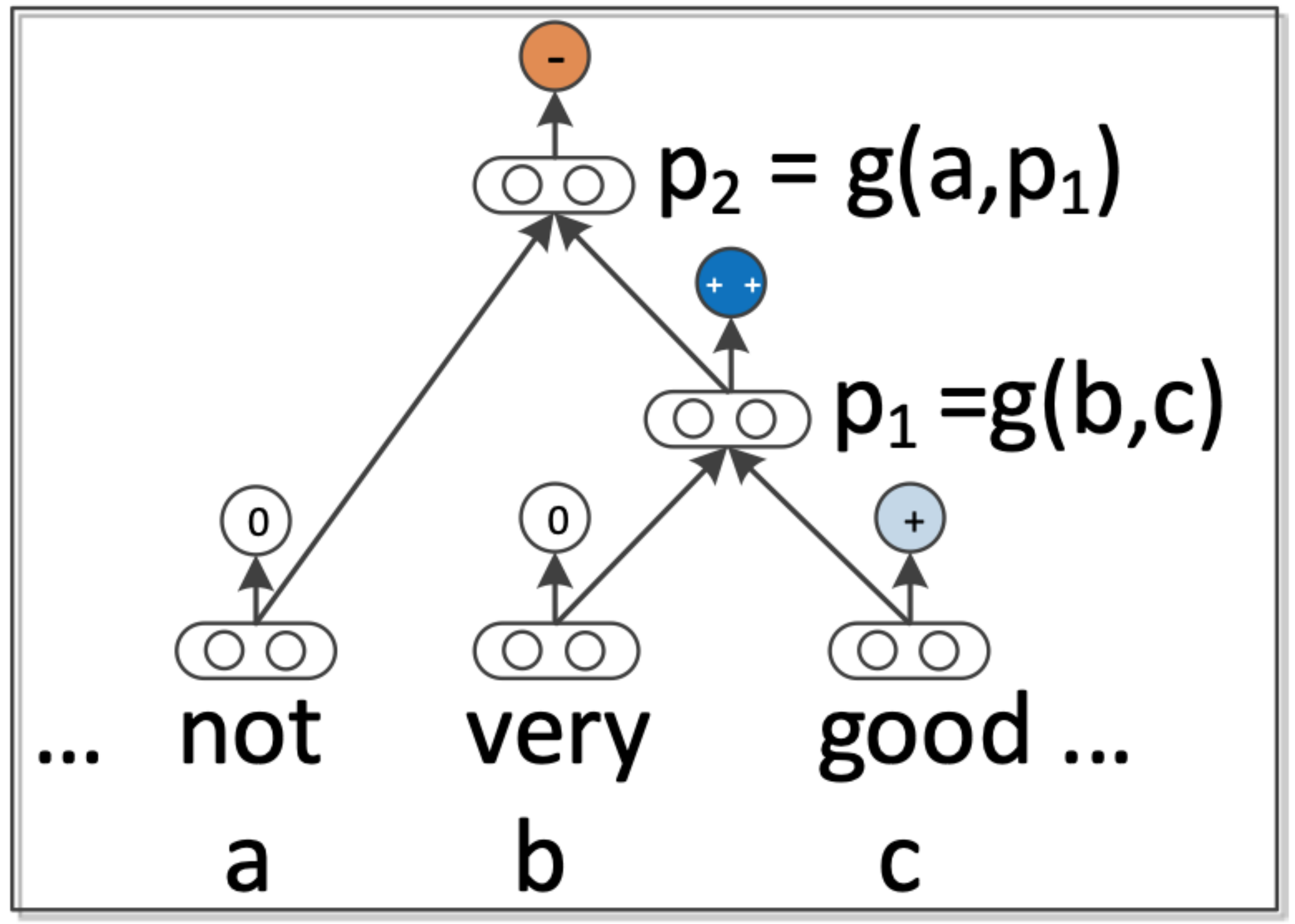
- Recurrent neural networks (Linear recursive neural networks)
- **Text as a sequence**
- Word embeddings + LSTM + Conditional random fields
- LSTM/Bi-LSTM capture long dependencies in sequences, learns what to remember and what to forget
- Named entity recognition, aspect level sentiment analysis



Deep learning architectures for text classification and sentiment analysis

- Recursive neural network: Same weights applied over a structured input
- **Text as a tree**, bottom-up representation





Recursive deep models for semantic compositionality over a sentiment treebank. Socher et al. (2013)

Deep learning architectures for text classification and sentiment analysis

- Convolutional neural networks
- **Finds local patterns in text**
- Applies filters/convolutions + pooling layers to find good features
- Softmax output layer for binary classification
- Input embeddings:
  - Use pre-trained embeddings
  - or initialize embeddings randomly and learn them when network is trained
  - or use pre-trained embeddings and fine-tune them during training

Deep learning architectures for text classification and sentiment analysis

- Transformer/Attention based language models : BERT et al.
- **Contextual representations of text**
- Pre-trained on very large text collections
- Very useful for text classification
- Different output layers can be suitable for different tasks

Which model is the best for text classification?

- Recent models in the BERT family perform very well and can be fine tuned for specific tasks and domains if some training data is available. New model variations with new state-of-the art results are reported continuously
- LAB: try classification with BERT
- Fine-tuning can be applied to only the output layers or include previous layers
- Different types of output layers can be appended to transformer models depending on the classification task
- Simpler models, such as Naive Bayes, can be used baselines
- Naive Bayes is fast, works well for small data sets and is easy to implement

Is sentiment analysis solved?

- Accuracy for binary sentiment classification of the IMDB dataset:
  - BERT large: 95.97%
  - XLNet large: 96.21%
- Both models have over 300 million parameters
- Training XLNet required 2000 GPU days
- Binary sentiment classification: Is this text mainly negative or positive?
- Still a long way to go before deeper language understanding is achieved!
- Remaining tasks and challenges:
  - Interpretability: what does the model understand/misunderstand? Why?
  - Environmental impact, can we keep building bigger models?
  - Bias in large language models: biased text as input will produce biased models

- Adam Bermingham and Alan Smeaton. On using twitter to monitor political sentiment and predict election results. In *Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2011)*, pages 2–10, 2011.
- Johan Bollen, Huina Mao, and Xiaojun Zeng. Twitter mood predicts the stock market. *Journal of computational science*, 2(1):1–8, 2011.
- Hai Ha Do, PWC Prasad, Angelika Maag, and Abeer Alsadoon. Deep learning for aspect-based sentiment analysis: a comparative review. *Expert Systems with Applications*, 118:272–299, 2019.
- Shervin Minaee, Nal Kalchbrenner, Erik Cambria, Narjes Nikzad, Meysam Chenaghlu, and Jianfeng Gao. Deep learning–based text classification: A comprehensive review. *ACM Computing Surveys (CSUR)*, 54(3):1–40, 2021.
- Saif Mohammad. Word affect intensities. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, May 2018. European Language Resources Association (ELRA).
- Lawrence Phillips, Chase Dowling, Kyle Shaffer, Nathan Hodas, and Svitlana Volkova. Using social media to predict the future: a systematic literature review. *arXiv preprint arXiv:1706.06134*, 2017.
- Andrew J Reagan, Lewis Mitchell, Dilan Kiley, Christopher M Danforth, and Peter Sheridan Dodds. The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5(1):1–12, 2016.
- Sonja Remmer, Anastasios Lamproudis, and Hercules Dalianis. Multi-label Diagnosis Classification of Swedish Discharge Summaries–ICD-10 Code Assignment Using KB-BERT. In *RANLP 2021: Recent Advances in Natural Language Processing, 1-3 Sept 2021, Varna, Bulgaria*, pages 1158–1166. Association for Computational Linguistics, 2021.
- Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D Manning, Andrew Y Ng, and Christopher Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceeding s of the 2013 conference on empirical methods in natural language processing*, pages 1631–1642, 2013.
- Ashima Yadav and Dinesh Kumar Vishwakarma. Sentiment analysis using deep learning architectures: a review. *Artificial Intelligence Review*, 53(6):4335–4385, 2020.

Thank you!