Sentiment Analysis

> ... *reviewing machine learning literature about sentiment analysis*

- the process of detecting polarity from positive to negative in text
- often used by businesses to detect sentiment in social data, gauge brand reputation, and understand customers
- types of sentiment analysis
    - fine-grained sentiment analysis : interpret polarity categories
    - emotion detection : aims to detect emotions (happiness, frustration, anger, sadness, etc.) and usually use lexicons (i.e. lists of words and the emotions they convey) or machine learning algorithms
    - aspect-based sentiment analysis : analysis of which particular aspects or features are mentioned i.e. The battery life of this camera is too short negative opinion about the feature battery life
    - multilingual sentiment analysis : difficult, need a lot of preprocessing and resources
- sentiment analysis algorithms
    - rule-based : these systems perform analysis based on a set of manually crafted rules
        - define two lists of polarised words, counts the number of positive and negative words that appear in a given text, system returns positive or negative by comparing the number
        - naive, don't take into account how words are combined in a sequence and don't support new vocabulary, and adding new rules may affect previous result
    - automatic : rely on machine learning techniques to learn from data
        - feature extraction from text : classical approach (bag-of-words or bag-of ngrams), recently word embeddings
        - classification algorithms : Naive Bayes, Linear Regression, Support Vector Machines, and Deep Learning
    - hybrid : systems combine both rules-based and automatic approaches
- challenges
    - subjectivity and tone
        - 'The package is nice.', 'The package is red.'
        - all predicates should not be treated the same with respect to how they create sentiment
    - context and polarity
        - 'Everything of it.', 'Absolutely nothing!'
        - responses of the question 'What did you like about the event?' vs 'What did you dislike about the event?'
        - pre processing or post processing needed to take into account context
    - irony and sarcasm
        - people express their negative sentiments using positive words
        - 'Did you enjoy your shopping experience with us?' 'Yeah, sure. So smooth!', 'Not one, but many!'
        - there is no textual cue that will help a machine learn
    - comparisons
        - how to treat comparisons
        - 'This product is second to none.', 'This is better than older tools.', 'This is better than nothing.'
    - emojis
        - play an important role in the sentiment of texts, particularly in tweets
    - defining neutral
        - what you mean by neutral, positive, or negative does matter when you train sentiment analysis models
        - i.e. include objective text which do not contain explicit sentiments, into the neutral category, irrelevant information tag as neutral etc.
    - human annotator accuracy
        - sentiment analysis is difficult task even for human
        - inter-annotator agreement (a measure of how well two (or more) human labellers can make the same annotation decision) is pretty low when it comes to sentiment analysis
- https://monkeylearn.com/sentiment-analysis/
- how we do aspect-based sentiment analysis
    - there are usually two steps
        1. extract aspect term
        2. sentiment analysis for each aspect
    - using NLP tools such as spaCy library, NLTK, word2vec, gensim etc.
    - example
        1. https://intellica-ai.medium.com/aspect-based-sentiment-analysis-everything-you-wanted-to-know-1be41572e238
        2. https://aclanthology.org/S14-2004/
            a. aspect term extraction - identify all aspect terms present in each sentence i.e. I like the *service* and *staff*, but not the *food.*
            b. aspect term polarity - find sentiment polarity from given aspect term i.e. I hated their fajitas, but their salads were great. {fajitas: neg., salads: pos.}
            c. aspect category detection - detect pre-defined aspect categories such as price or food i.e. The restaurant was expensive, but the menu was great." {price, food}
            d. aspect category polarity

Multiple Instance Learning Networks for Fine-Grained Sentiment Analysis

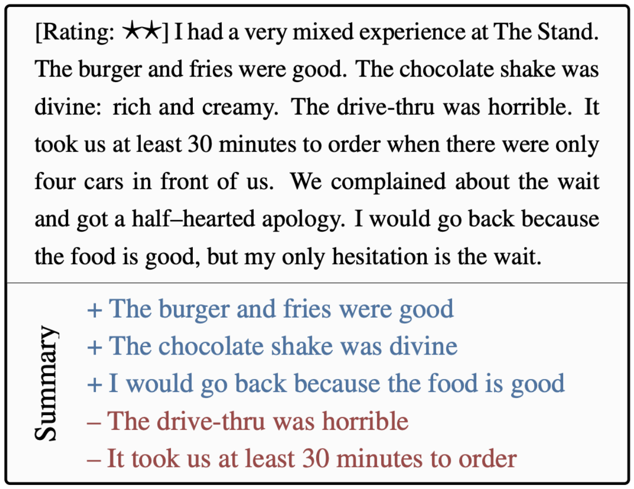
- Goal : we consider the problem of segment level sentiment analysis from the perspective of Multiple Instance Learning
- Link : https://aclanthology.org/Q18-1002/ (2018)
- Code : https://github.com/stangelid/oposum
- Credible source : Transactions of the Association for Computational Linguistics (TACL)

Model

<table>
<tr><th>approaches</th><th>description</th></tr>
<tr><td>

- document level label
  - coarse-grained
  - easy to obtain due to the widespread use of opinion grading interfaces
- sentence or phrase level label
  - finer-grained
  - laborious and expensive work
  - as a whole, the review conveys negative sentiment, aspects of the reviewer's experience were clearly positive
  - these positive sentences are unnoticed when focusing solely on the review's overall rating

</td><td>



</td></tr>
<tr><td>

- Elementary Discourse Units (EDUs)
  - adopted from Rhetorical Structure Theory's (Mann and Thompson, 1988)
    - According to RST, documents are first segmented into EDUs corresponding roughly to independent clauses which are then recursively combined into larger discourse spans.
    - This results in a tree representation of the document, where connected nodes are characterised by discourse relations.
    - We only utilise RST's segmentation, and leave the potential use of the tree structure to future work.
  - definitions for EDUs vary in the literature, we follow standard practice and take the elementary units of discourse to be clauses (Carlson et al., 2003)
  - employ a state-of-the-art discourse parser (Feng and Hirst, 2012) to identify them

</td><td>

- EDU based segmentation might be beneficial for opinion extraction.
- The second and third EDUs correspond to the sentence: I didn't enjoy most of them, but the burger was brilliant. Taken as a whole, the sentence conveys mixed sentiment, whereas the EDUs clearly convey opposing sentiment.

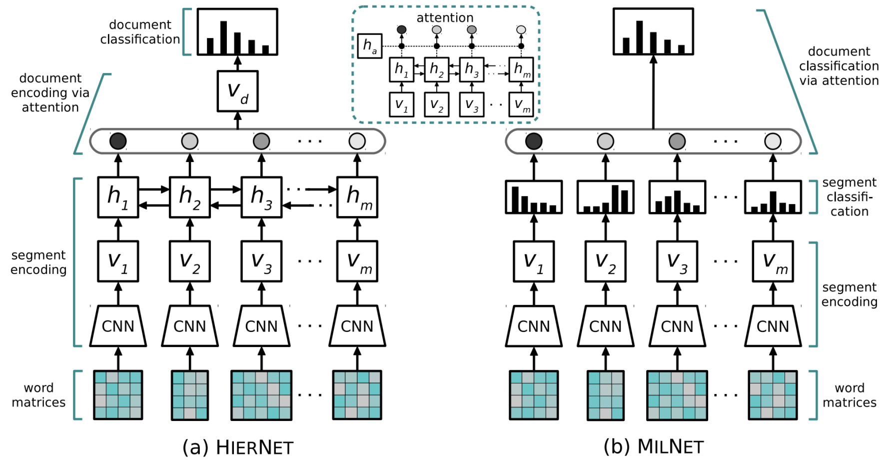
</td></tr>
</table>

- Sentimental classification
  - creation of sentimental lexicons based on which the overall polarity of a text
  - SO-CAL, a state-of-the-art method that combines a rich sentiment lexicon with carefully defined rules over syntax trees
  - neural network models
  - CNN architecture for sentence-level classification
- Hierarchical Network (HIERNET)
  - building representations of sentences and aggregating those into a document feature vector
  - given document d comprising segments s1, s2, ... sm
  - using CNN, produce segment representations v1, v2, ... vm and hidden vectors h1, h2, ... hm
  - hidden vectors are used to produce attention weight a1, a2, ... am
  - document representation Vd is the weighted average of the segments' hidden vectors

    $$\mathrm{v}_d = \sum_i a_i \mathrm{h}_i.$$

  - final sentiment prediction is obtained using a softmax classifier
  - predict document-level polarity by encoding sentences and then combining these representations into a document vector

- Multiple Instance Learning (MIL)
  - deals with problems where labels are associated with groups of instances or bags (documents in our case), while instance labels (segment-level polarities) are unobserved
  - an aggregation function is used to combine instance predictions and assign labels on the bag level
  - the goal is either to label bags or to simultaneously infer bag and instance labels
  - based on the assumption
    - each segment conveys a degree of sentiment polarity, ranging from very negative to very positive
    - segments have varying degrees of importance, in relation to the overall opinion of the author
    - the overarching polarity of a text is an aggregation of segment polarities, weighted by their importance
    - model attempts to predict the polarity of segments and decides which parts of the document are good indicators of its overall sentiment



- Multiple Instance Learning Network (MILNET)
  - segment encoding : an encoding is produced for each segment using CNN
  - segment classification : achieve individual distributions using softmax from seperate representation vi
  - document classification : document-level predictions can be produced by taking the average of segment class distributions

- polarity-based opinion extraction
  - after training, model can produce segment-level sentiment predictions for unseen text in the form of class probability distributions
  - polarity score : compute the polarity score of a segment as the dot-product of the probability distribution pi with vector w:

$$\text{polarity}(s_i) = \sum_c p_i^{(c)} w^{(c)} \quad \in [-1, 1]$$

  - gated polarity : uses the attention mechanism to differentiate between segments that carry significant sentiment cues and those that do not:

$$\text{gated-polarity}(s_i) = a_i \cdot \text{polarity}(s_i)$$

Experiment

training data	experiment	evaluation	output

training data

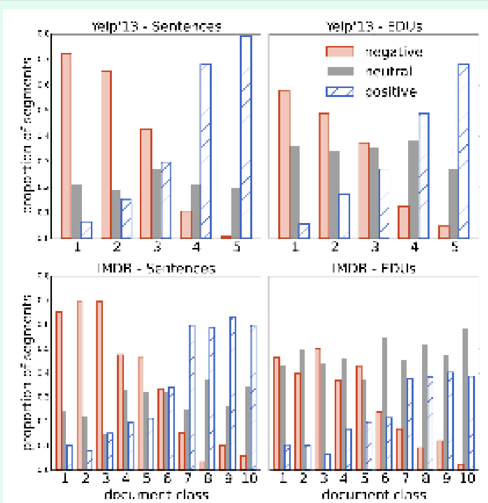
two large-scale sentiment classification collections

experiment

- train MILNET model, HIERNET model
  - 25 epochs
  - mini-batches of 200 documents,

evaluation

- to evaluate model performance, construct SPOT (Segment-level POlariTy) dataset
  - sample reviews from each collections
  - such that all document-level classes are represented uniformly
  - and the document lengths are representative of the respective corpus
  - documents are segmented into two segment-level datasets, sentences and EDUs
- each review was presented to three Amazon Mechanical Turk (AMT) annotators, and annotated as negative, neutral, or positive
- Amazon Mechanical Turk (AMT)
  - a crowdsourcing website for businesses to hire remotely crowd workers to perform discrete on demand tasks that computers are currently unable to do
  - https://en.wikipedia.org/wiki/Amazon_Mechanical_Turk
  - employers post jobs known as Human Intelligence Tasks (HITs), workers complete them in exchange for a rate set by the employer

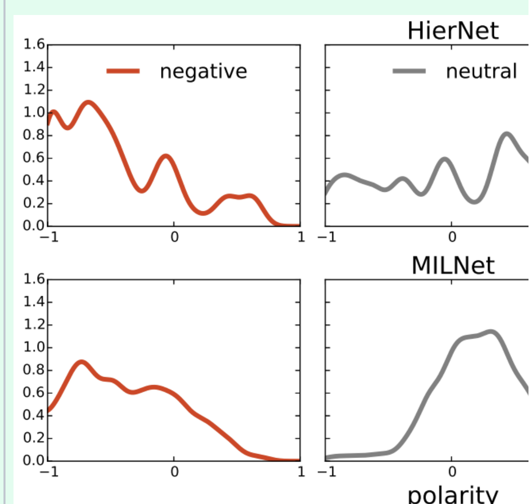


- distribution of segment labels for each document-level class

output

	Method	Yelp'13$_{seg}$ Sent	Yelp'13$_{seg}$ EDU	IMDB$_{seg}$ Sent	IMDB$_{seg}$ EDU
	Majority	19.02†	17.03†	18.32†	21.52†
Document	HierNet$_{avg}$	54.21†	50.90†	46.99†	49.02†
Document	HierNet	55.33†	51.43†	48.47†	49.70†
Document	HierNet$_{gt}$	56.64†	58.75	62.12	57.38†
Document	MilNet$_{avg}$	58.43†	48.63†	53.40†	51.81†
Document	MilNet	52.73†	53.59†	48.75†	47.18†
Document	MilNet$_{gt}$	59.74†	59.47	61.83†	58.24†
Segm	MilNet$_{avg}$	51.79†	46.77†	45.69†	38.37†
Segm	MilNet	61.41	59.58	59.99†	57.71†
Segm	MilNet$_{gt}$	**63.35**	**59.85**	**63.97**	**59.87**
	SO-CAL	56.53†	58.16†	53.21†	60.40
	Seg-CNN	56.18†	59.96	58.32†	62.95†

- second block : models that do not utilise segment-level pred segments using its document-level predictions
- third block : segment-level predictions
- three levels of attention integration : model without gating(n model trained with the attention mechanism disabled(avg s
- MILNET with gated, segment specific polarities obtains the



- distribution of polarity scores produced by the two models c
- neutral class appears to be problematic for HIERNET, polar range of values, while the MILNET's distribution is near zer

Experimental results demonstrate the superior performance of o architectures

1. The Yelp 13 corpus a. contains customer reviews of local businesses b. each associated with human ratings on a scale from 1 to 5	• 300 dimensional pre-trained word2vec embedding • window size 3, 4 in CNN segment encoder		• documents with positive labels contain a large number of positive segments compared to negative labelled document	

2. The IMDB corpus a. contains movie reviews b. each review is associated with user ratings ranging from 1 to 10 both are split into training (80%), validation (10%), and test(10%) sets	• 5 words with 100 feature maps per window size • hidden vector dimensions for each direction is 50		• neutral segments are distributed in an approximately uniform manner across document classes • the proportion of neutral EDU	

	• attention vector dimensionality is 100 • L2-normalisation • dropout to regularise the softmax classifier		s is significantly higher than compared to neutral sentences the observation reinforces our argument in favour of EDUs	

	• compare MILNET against following methods (baseline) : Majority, SO-CAL, Seg-CNN, GICF, HIERNET + produce finer-grained polarity distinctions via gating, using the model's attention weights		egmentation, as it suggests that a sentence with positive or negative overall polarity may still contain neutral EDUs	

- discarding neutral EDUs, could therefore lead to more concise opinion extraction compared to relying on entire sentences

		• macro-averaged F1
		  ◦ used to assess the quality of problems with multiple binary labels or multiple classes
		  ◦ defined as the mean of class-wise/label-wise
		
		• evaluate 2 points
		  1. assess models' ability to classify segment polarity in reviews using the newly created SPOT dataset
		  2. focused on opinion extraction: conduct a judgment elicitation study to determine whether extracts produced by MILNET are useful and of higher quality compared to HIERNET and other baselines	

Data

bags_and_cases.trn

```
B00001OPGT-0000 2
This case looks nice EDU_BREAK and has plenty of pockets for all the other stuff EDU_BREAK you have to car
But , EDU_BREAK if you 're using it for something other than back and forth to the office or have a thinne
just not enough padding in here EDU_BREAK to protect the computer from scratching and rubbing .
We bought this for my son for school this past Fall .
He has one of those thinner laptops EDU_BREAK and the cover was soon dented and scratched EDU_BREAK and th
screen on the touchpad controls left little white marks on the display .
While you can blame the computer design some , EDU_BREAK after we added more padding in the case , EDU_BRE
On the plus side , EDU_BREAK the case is durable - EDU_BREAK nothing has yet ripped or torn off - EDU_BREA
clean .

B00001OPGT-0001 3
I bought this one few months back for my Toshiba laptop ... . . the case is quite strong EDU_BREAK thou a
Toshiba Satellite ... EDU_BREAK but on the other side this bag is quite strong EDU_BREAK and sturdy ... ED
EDU_BREAK if u r EDU_BREAK thinking EDU_BREAK that u can also put ur files EDU_BREAK n other documents in
laptop then this is not the bag for u. ... Pros :
Very Strong , Good Quality fabric n zippers , Water Proof , Very Light and just the right size , Smart loo
for 15 '' Laptops , wont carry ur books or big files ... EDU_BREAK overall i think this a very good bag fo

B00001OPGT-0002 3
The pic on this product does no justice to the real thing .
Its very good looking EDU_BREAK -LRB- BLACK synthetic with a sheen -RRB- , EDU_BREAK and well designed .
Isnt one of those huge contraptions EDU_BREAK that people carry .
The hand and shoulder straps have delicious padding !! Adequate space for a large computer EDU_BREAK -LRB-
EDU_BREAK a zip drive , cables , some disks , magazines , notes etc. .
If your looking to carry bulky files and text books - EDU_BREAK might not work .
Value for money any day .
```

Repository

Sentiment Analysis Based on Deep Learning: A Comparative Study

- Goal : reviews the latest studies that have employed deep learning to solve sentiment analysis problems, such as sentiment polarity
- Link : https://arxiv.org/abs/2006.03541 (06/2020)
- Code : -
- Credible source : -

A Unified Generative Framework for Aspect-Based Sentiment Analysis

- Goal : redefine every subtask target as a sequence mixed by pointer indexes and sentiment class indexes, which converts all ABSA subtasks into a unified generative formulation, exploiting the pre-training sequence-to-sequence model BART to solve all ABSA subtasks in an end-to-end framework
- Link : https://arxiv.org/abs/2106.04300 (06/2021)
- Code : https://github.com/yhcc/BARTABSA
- Credible source: -