From opinion lexicons to sentiment classification of tweets and vice versa: a transfer learning approach

2016 IEEE/WIC/ACM International Conference on Web Intelligence
Omaha, Nebraska, USA

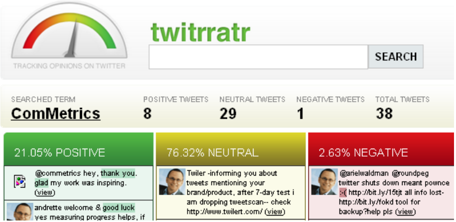
Felipe Bravo-Marquez, Eibe Frank, and Bernhard Pfahringer

Department of Computer Science, University of Waikato

16 October, 2016

Sentiment Analysis and Social Media

- Twitter users tend to publish **personal opinions** regarding certain topics and news events.
    1. Hey @Apple, pretty much all your products are amazing. You blow minds every time you launch a new gizmo. That said, your hold music is crap.
    2. #windows sucks... I want #imac so bad!!! why is it so damn expensive :(@apple please give me free imac and I will love you :D
- Analysing the sentiment underlying these opinions has important applications in product **marketing** and **politics**.



Opinion Mining or Sentiment Analysis

Twitter Sentiment Analysis Main Tasks

1. **Message-level polarity classification** (MPC): classify tweets into sentiment categories such as **positive** and **negative**.
2. **Polarity lexicon induction** (PLI): classify words from a corpus of tweets into sentiment categories, e.g., **happy, wonderful**, **sad, bad**.

- State-of-the-art solutions use **supervised** machine learning models trained from **manually** annotated examples.
- **Problem**: annotation of words or tweets based on polarity classes is a **time-consuming** and **labor-intensive task**.
- **Possible Solution**: Transferring **existing labels** from a related problem domain.

Transfer Learning Approach

- **Transfer learning**: improving learning task for a **target domain** $\mathcal{D}_{\mathcal{T}}$ using knowledge obtained from a related **source domain** $\mathcal{D}_{\mathcal{S}}$.
- We transfer **sentiment knowledge** from the word domain $\mathcal{D}_{\mathcal{W}}$ to the message domain $\mathcal{D}_{\mathcal{M}}$ and **vice versa**.
- Tweets and words can be labelled according to the **same sentiment categories**, e.g, positive and negative ($\mathcal{Y}_{\mathcal{W}} = \mathcal{Y}_{\mathcal{M}}$).
- We propose a **unified representation** that allows the bidirectional transfer of sentiment classifiers between words and tweets.

The word-tweet sentiment-interdependence relation

1. The polarity of a tweet is **determined** by the polarity of the words it **contains**.
2. The polarity of a word is **determined** by the polarity of the tweets in which it **occurs**.

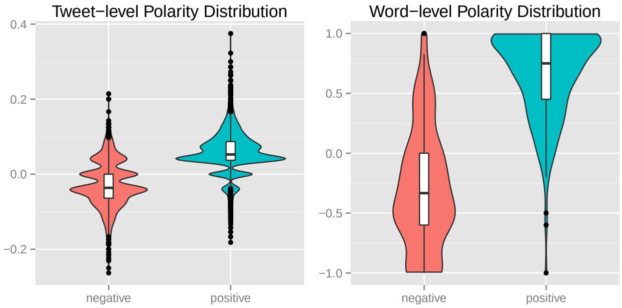


Figure : Violin plots of the polarity of tweets and words.

Transfer Learning with Tweet Centroids

- We represent tweets and words by **feature vectors** of the **same dimensionality**.
- Tweets are represented using **NLP features**: 1) unigrams, 2) part-of-speech (POS) tags, and 3) Brown words clusters.
- **Tweet Centroid Model (TCM)**: words are represented by the **centroids** of the tweet vectors in which they **occur**.
- TCM allows **classifiers** trained from one of the two domains to be **deployed** on data from the other.

Transfer Learning with Tweet Centroids (2)

vocabulary

w1	angry
w2	happy
w3	good
w4	grr
w5	lol

tweets

lol happy
lol good
grr angry

Word vectors

	w1	w2	w3	w4	w5
w1	1	0	0	1	0
w2	0	1	0	0	1
w3	0	0	1	0	1
w4	1	0	0	1	0
w5	0	0.5	0.5	0	1

$$\sum_{t \in \mathcal{M}(w)} \frac{x_j^{(t)}}{|\mathcal{M}(w)|}$$

Tweet centroids

Tweet vectors

	w1	w2	w3	w4	w5
t1	0	1	0	0	1
t2	0	0	1	0	1
t3	1	0	0	1	0

Label words and train word-level classifier

Label tweets and train tweet-level classifier

Target words

lol	pos
grr	neg

Target tweets

Happy morning	pos
What a bummer!	neg
Lovely day	pos

Inducing a Lexicon from Labelled Tweets

- We can train a **message-level classifier** f_M from a corpus of sentiment annotated tweets $\mathcal{C}_L$ (SemEval) and deploy it on words found in a **corpus of unlabelled tweets** $\mathcal{C}_U$ represented by tweet centroids.
- We calculate the word-vectors from a **larger** corpus of unlabelled tweets (2M) to get **better** representations.

AUC		
Source Dataset	PMI-SO	TCM
Sanders	0.757	**0.864**
6HumanCoded	0.861	**0.930**
SemEval	0.858	**0.916**

Table : Word-level Polarity Classification Results for the AFINN lexicon.

- TCM **outperformed** PMI-SO, a state-of-the-art measure for establishing world-level sentiment.

Inducing a Lexicon from Labelled Tweets

Figure : Word clouds of positive and negative words obtained from a message-level classifier.

Tweet Centroids for message-level classification

- TCM can be used as a **distant supervision** model for MPC.
- We use a **word-level** classifier f_W trained with TCM vectors calculated from $\mathcal{C}_U$ labelled by a **polarity lexicon** $\mathcal{L}$.
- The classifier is deployed on the target tweets represented by **sparse vectors**.
- The number of labelled words for training f_W is **limited** to the number of words from $\mathcal{L}$.
- TCM is **not capable** of exploiting large collections of unlabelled tweets for producing training datasets larger than the size of $\mathcal{L}$.

Partitioned TCM

- We propose a modification of our method for **increasing** the number labelled instances it produces.
- The word-tweet set $\mathcal{M}(w)$ for each word from the lexicon ($w \in \mathcal{L}$) is **partitioned** into smaller disjoint subsets $\mathcal{M}(w)_1, \ldots \mathcal{M}(w)_z$ of a fixed size determined by a parameter p.
- We calculate one tweet centroid vector $\overrightarrow{w}$ for **each partition** labelled according to $\mathcal{L}$.

Baselines

Emoticon-Annotation Approach (EAA)

- Labels tweets with positive or negative emoticons according to the emoticon's polarity after removing the emoticon from the message.
- Tweets containing both positive and negative emoticons are **discarded**.

Lexicon-annotation approach (LAA)

- Uses a given polarity lexicon $\mathcal{L}$.
- Tweets with at least one positive word and no negative word are labelled **positive**.
- Tweets with at least one negative word and no positive word are labelled **negative**.

Instances Generated by Distant Supervision Models

We use 10 collections of 2 million tweets as source corpora.

	Avg. Positive	(%)	Avg. Negative	(%)	Avg. Total	(%)
EAA	130,641	(6.5%)	21,537	(1.1%)	152,179	(7.6%)
LAA	681,531	(34.1%)	294,177	(14.7%)	975,708	(48.8%)
TCM	1537	(0.05%)	951	(0.08%)	2488	(0.12%)
TCM (p=5)	276,696	(13.8%)	149,989	(7.5%)	426,684	(21.3%)
TCM (p=10)	138,596	(6.9%)	75,390	(3.8%)	213,986	(10.7%)
TCM (p=20)	69,518	(3.5%)	38,044	(1.9%)	107,563	(5.4%)
TCM (p=50)	32,231	(1.6%)	17,950	(0.9%)	50,181	(2.5%)
TCM (p=100)	14,338	(0.7%)	8357	(0.4%)	22,695	(1.1%)

TCM for MPC

	6HumanCoded		Sanders		SemEval	
EAA	0.805 ± 0.005	= -	0.800 ± 0.017	= +	0.802 ± 0.006	= -
LAA	0.809 ± 0.001	+ =	0.778 ± 0.002	- =	0.814 ± 0.000	+ =
TCM	0.776 ± 0.004	- -	0.682 ± 0.024	- -	0.779 ± 0.008	- -
TCM (p=5)	0.834 ± 0.002	+ +	0.807 ± 0.008	= +	0.833 ± 0.002	+ +
TCM (p=10)	0.845 ± 0.003	+ +	**0.817** ± 0.006	+ +	0.841 ± 0.002	+ +
TCM (p=20)	**0.850** ± 0.003	+ +	0.815 ± 0.011	+ +	**0.844** ± 0.003	+ +
TCM (p=50)	0.844 ± 0.004	+ +	0.785 ± 0.010	- +	0.836 ± 0.004	+ +
TCM (p=100)	0.829 ± 0.003	+ +	0.752 ± 0.019	- -	0.821 ± 0.004	+ +

Table : AUC for Message-level Polarity Classification. Best results per column are given in bold.

Conclusions

- We proposed a **distant supervision method** that outperformed LAA and EAA for MPC.
- TCM is a **unified model** for message-level and word-level sentiment classification.
- Future work: subjectivity, emotions, handle negations, non-linear representations and deep networks.

Questions?

Thanks for your Attention!

Acknowledgements

- University of Waikato Doctoral Scholarship
- Machine Learning Group at the University of Waikato

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

References I

Kiritchenko, S., Zhu, X., and Mohammad, S. M. (2014).
Sentiment analysis of short informal texts.
Journal of Artificial Intelligence Research, 50:723–762.

Mintz, M., Bills, S., Snow, R., and Jurafsky, D. (2009).
Distant supervision for relation extraction without labeled data.
In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2 - Volume 2*, ACL '09, pages 1003–1011, Stroudsburg, PA, USA. Association for Computational Linguistics.

Read, J. (2005).
Using emoticons to reduce dependency in machine learning techniques for sentiment classification.
In *Proceedings of the ACL Student Research Workshop*, ACLstudent '05, pages 43–48, Stroudsburg, PA, USA. Association for Computational Linguistics.