Positive, Negative, or Neutral: Learning an Expanded Opinion Lexicon from Emoticon-annotated Tweets

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

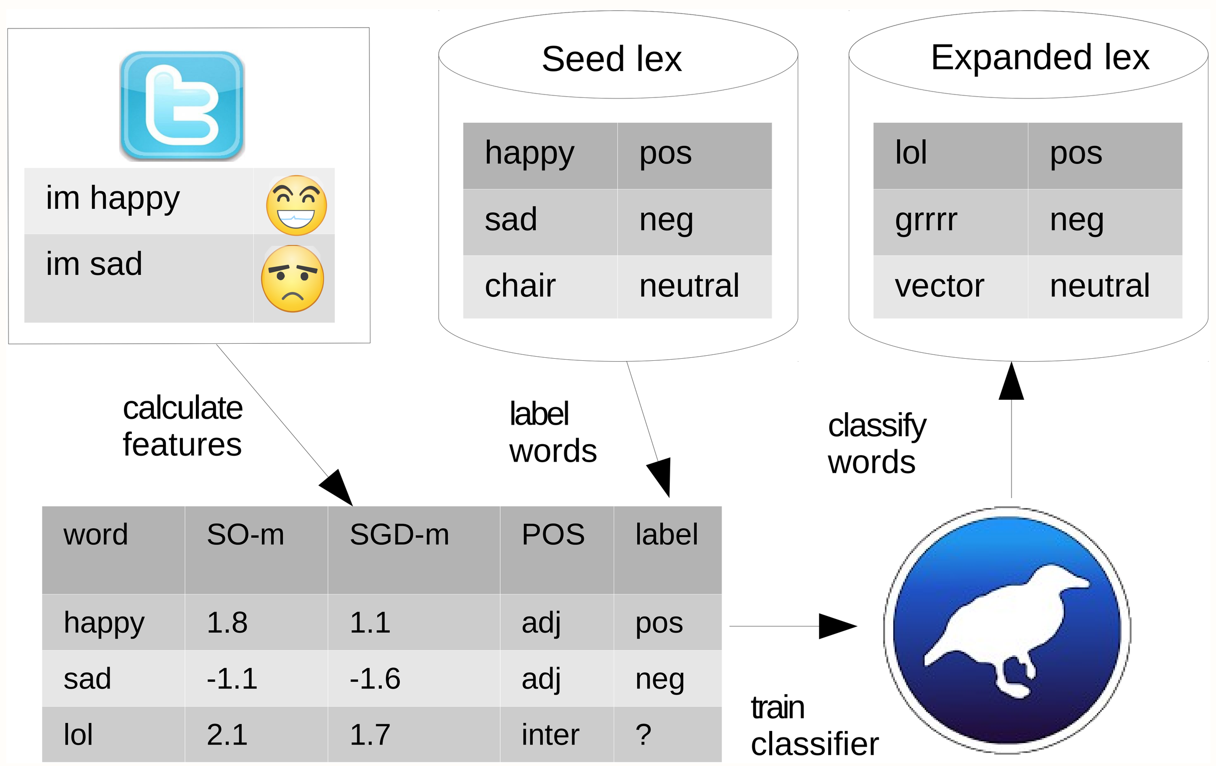
Felipe Bravo-Marquez, Eibe Frank Bernhard Pfahringer
Department of Computer Science, University of Waikato

Problem

- An opinion lexicon is a lists of terms labelled by sentiment.
- They are normally composed of positive and negative words such as **happy, wonderful** and **sad, bad**.
- The words used in Twitter include many abbreviations, acronyms, and misspelled words, e.g., **lol**, **omg**, **hahaha**, **#hatemonday** that are **not** covered by most popular lexicons.
- The manual creation of a Twitter-oriented opinion lexicon is a **time-consuming** task.

Solution

- We propose a **supervised model** for **Twitter** lexicon expansion from **emoticon** annotated tweets and a **seed lexicon**.
- We propose word-level attributes based on part-of-speech tags (**POS**), stochastic gradient descent (**SGD**), and semantic orientation (**SO**).
- The lexicon contains POS disambiguated entries with a three-dimensional probability distribution for **positive**, **negative**, and **neutral** polarities.



Ground-Truth word polarities

- To label the words we create a **seed** lexicon by taking the **union** of existing hand-made lexicons and discarding all words where a **polarity clash** is observed.

	Positive	Negative	Neutral
AFINN	564	964	0
Bing Liu	2003	4782	0
MPQA	2295	4148	424
NRC-Emo	2312	3324	7714
Seed Lex	3730	6368	7088

Obtaining Emoticon-annotated Tweets

- We consider **two** collections of tweets covering multiple topics: The **Edinburgh corpus** (ED), and the **Stanford Sentiment corpus** (STS).
- Tweets exhibiting **positive :)** and **negative :(** emoticons are labelled according to the emoticon's polarity.

	ED	STS
Positive	1,813,705	800,000
Negative	324,917	800,000
Total	2,138,622	1,600,000

Word-level features

- Tweets are lowercased, tokenised and POS-tagged.
- We include the POS-tag of the word as a nominal attribute.
- We also create two **time-series** for each word: the **Stochastic Gradient Descent** (SGD) series, and the **Semantic Orientation** (SO) series, from which we extract additional features.

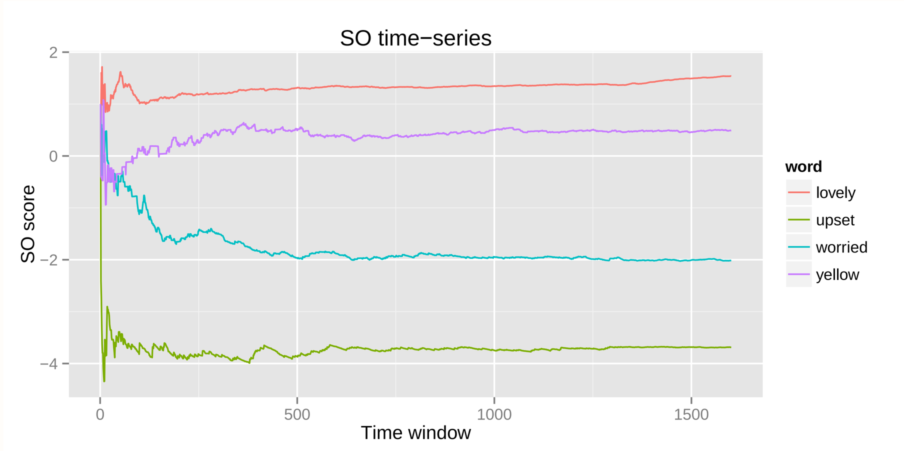
Feature	Description
mean	The mean of the time-series.
trunc.mean	The truncated mean of the time-series.
median	The median of the time-series.
last.element	The last observation of the time-series.
sd	The standard deviation of the time-series .
iqr	The inter-quartile range.
sg	The fraction of times the time-series changes its sign.
sg.diff	The sg value for the differenced time-series.

- The SGD series is calculated by incrementally training a **linear support vector machine** from the stream of labelled tweets.
- We use **stochastic gradient descent** (SGD) online learning process.

$$\frac{\lambda}{2}||w||^2 + \sum [1 - y(\mathbf{x}\mathbf{w} + b)]_+.$$

- The weights of this linear model correspond to POS-tagged words and are updated in an **incremental fashion**.
- The model's weights determine how strongly the presence of a word **influences** the prediction of **polarity** classes.
- The SO series corresponds to the **accumulated semantic orientation** (SO) [2] and is based on the **point-wise mutual information** measure.

$$\text{SO} = \log_2\left(\frac{\text{count}(\text{word} \wedge \text{pos}) \times \text{count}(\text{neg})}{\text{count}(\text{word} \wedge \text{neg}) \times \text{count}(\text{pos})}\right)$$



Experiments

- We study **three word-level** classification problems.
- **Neutrality**: Classify words as neutral (objective) or non-neutral (subjective).
- **PosNeg**: Classify words to positive or negative classes.
- **Polarity**: Classify words to classes positive, negative or neutral. This is the classification problem we aim to solve.
- We trained **RBF SVM classifiers** for the different problems in both datasets.
- We compare the weighted **AUC** obtained by classifier based solely on **SO** with a classifier that uses all the features.

Dataset	SO	ALL
ED-Neutrality	0.62 $\pm$0.02	**0.65**$\pm$0.02$\circ$
ED-PosNeg	0.74 $\pm$0.03	**0.75**$\pm$0.03
ED-Polarity	0.62 $\pm$0.02	**0.65**$\pm$0.02$\circ$
STS-Neutrality	0.63 $\pm$0.02	**0.67**$\pm$0.02$\circ$
STS-PosNeg	**0.77**$\pm$0.03	**0.77**$\pm$0.03
STS-Polarity	0.64 $\pm$0.02	**0.66**$\pm$0.01$\circ$

- We fit a logistic regression model to the output of the support vector machine trained for the *polarity* problem to classify the **unlabelled** words and create the **expanded** lexicon.

word	POS	negative	neutral	positive
alrighty	interjection	0.021	0.087	0.892
boooooo	interjection	0.984	0.013	0.003
lmaoo	interjection	0.19	0.338	0.472
french	adjective	0.357	0.358	0.285
handsome	adjective	0.007	0.026	0.968
saddest	adjective	0.998	0.002	0
same	adjective	0.604	0.195	0.201
anniversary	common.noun	0.074	0.586	0.339
tear	common.noun	0.833	0.124	0.044
relaxing	verb	0.064	0.244	0.692
wikipedia	proper.noun	0.102	0.644	0.254

- We use the provided probabilities to explore the sentiment **intensities** of words.

Positive intensities Negative intensities

- The sizes of the words are proportional to the **log odds ratios** $\log_2(\frac{P(\text{pos})}{P(\text{neg})})$ and $\log_2(\frac{P(\text{neg})}{P(\text{pos})})$ for positive and negative words, respectively.
- We compare the expanded lexicons based on ED and STS with the seed lexicon for categorising entire tweets into **positive** or **negative** sentiment classes.

Dataset	Baseline	ED	STS
6-coded	0.77 $\pm$ 0.03	**0.82** $\pm$ 0.03 $\circ$	**0.82** $\pm$ 0.02 $\circ$
Sanders	0.77 $\pm$ 0.04	0.83 $\pm$ 0.04 $\circ$	**0.84** $\pm$ 0.04 $\circ$
SemEval	0.77 $\pm$ 0.02	0.81 $\pm$ 0.02 $\circ$	**0.83** $\pm$ 0.02 $\circ$

Conclusions

- The method creates a lexicon with **disambiguated** POS entries and a probability distribution for **positive, negative, and neutral** classes.
- The method outperforms the three-dimensional word-level polarity classification performance obtained by semantic orientation [2].
- Sentiment analysis methods that are based on **SentiWordnet** [1] can be easily adapted to **Twitter** by relying on our lexicon.
- This method could be used to create **domain-specific** lexicons.
- It could also be used to study the **dynamics** of opinion-words.

References

[1] Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. LREC'10, pages 2200–2204, Valletta, Malta, 2010.

[2] Peter D. Turney. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. ACL '02, pages 417–424, Stroudsburg, PA, USA, 2002.