WEI CHEN

Oct. 8th

Origin of data:

Voice of customers

News

Tasks:

Polarity detection: whether the expressed opinion in a document, a sentence or an entity feature/aspect is positive, negative, or neutral.

Emotion classification: "beyond polarity" sentiment classification looks, for instance, at emotional states such as "angry", "sad", and "happy".

feature-based level: determining the opinions or sentiments expressed on different features or aspects of entities.

semi-supervised: make use of unlabelled data for training – typically a small amount of labeled data with a large amount of unlabelled data.

Methods:

knowledge-based techniques, based on the presence of unambiguous affect words such as happy, sad, afraid, and bored

statistical methods, latent semantic analysis, support vector machines, "bag of words" and *Semantic Orientation*

hybrid approaches.[both machine learning and elements from knowledge representation such as ontologies and semantic networks in order to detect semantics that are expressed in a subtle manner

WEI CHEN

Oct.15th

Active Deep Networks for Semi-Supervised Sentiment Classification

http://www.aclweb.org/anthology/C/C10/C10-2173.pdf

- Data set:

MOV, movie reviews from IMDb. http://www.cs.cornell.edu/people/pabo/movie-review-data/

BOO, DVD, ELE, KIT (product reviews from Amazon)

http://www.cs.jhu.edu/~mdredze/datasets/sentiment/

- Model input/output:

Each review is represented as a vector of unigrams: 1 for terms present in the reviews, 0 for absent terms.

Sort the dictionary by document frequency and remove the top 1.5% (stopwords or domain specific general-purpose words).

$y \in \{-1,1\}$(polarity detection)

- Active learning:

Initially we have L labeled training data.

for i=1 to I

Carry out the supervised learning with all labeled training data, based on active deep networks.

$$h_t^k(\mathbf{x}) = \mathrm{sigm}(c_t^k + \sum_{s=1}^{D_{k-1}} w_{st}^k h_s^{k-1}(\mathbf{x})) \qquad t = 1,\ldots,D_k$$

$$k = 1,\ldots,N-1$$

$$\arg\min_{\mathrm{h}^N} \quad \mathrm{f}\left(\mathrm{h}^N\left(\mathbf{X}^L\right),\mathbf{Y}^L\right)$$

where

$$\mathrm{f}\left(\mathrm{h}^N\left(\mathbf{X}^L\right),\mathbf{Y}^L\right) = \sum_{i=1}^{L}\sum_{j=1}^{C} \mathrm{T}\left(\mathrm{h}_j^N\left(\mathbf{x}^i\right) y_j^i\right)$$

and the loss function is defined as

$$\mathrm{T}(r) = \exp(-r)$$

Select G unlabeled reviews that are nearest to the separating hyperplane, and label them manually. **(main idea of active learning)**

end for

Reviews in this field

A review paper in 2012 which introduces every recent development of sentiment analysis in detail: Cross-Domain Sentiment Classification, Cross-Language Sentiment Classification, Dealing with Conditional Sentences, Sarcastic Sentences, Aspect-based Sentiment Analysis, Sentiment Lexicon Generation, Spam Detection, etc.

https://www.cs.uic.edu/~liub/FBS/liub-SA-and-OM-book.pdf

Review paper in 2014, an overview of the last update during 2010-2013, including many recently proposed algorithms, enhancements and various SA applications.

http://www.sciencedirect.com/science/article/pii/S2090447914000550

Survey on Aspect-Level Sentiment Analysis (2016)

https://personal.eur.nl/frasincar/papers/TKDE2016/tkde2016.pdf

A very recent paper classifies the same movie review dataset based on feature selection (2017) that may provide the state of the art benchmark of this domain. **(around 85% for polarity detection)**

https://www.waset.org/downloads/16/papers/17nl010049.pdf

TABLE IV Accuracy (%) of different classifiers with different feature size when proposed hybrid feature selection is used.

Features	NB	SVM	Adaboost	Bagging
1000	81	85.5	66.8	79.1
2000	85.4	87.2	66.7	80.1
3000	86.2	88.9	68.8	79.2
4000	85.5	88.1	67.2	79.5

SemEval (**Sem**antic **Eval**uation) is an ongoing series of evaluations of computational semantic analysis systems since 2007.

This yearly evaluation focus on Sentiment Analysis in Twitter for many years.

The task 4 in 2017: **Sentiment Analysis in Twitter**

http://alt.qcri.org/semeval2017/task4/

Task description:

Subtask A. (rerun): Message Polarity Classification: Given a message, classify whether the message is of positive, negative, or neutral sentiment.

Subtasks B-C. (rerun): Topic-Based Message Polarity Classification:

Given a message and a topic, classify the message on

B) two-point scale: positive or negative sentiment towards that topic

C) five-point scale: sentiment conveyed by that tweet towards the topic on a five-point scale.

Subtasks D-E. (rerun): Tweet quantification:

Given a set of tweets about a given topic, estimate the distribution of the tweets across

D) two-point scale: the "Positive" and "Negative" classes

E) five-point scale: the five classes of a five-point scale.

Dataset: Twitter messages on a range of topics in English and Arabic.

Training data in English: https://www.dropbox.com/s/qqvokdtalf0kgs2/2017_English_final.zip

Survey of the submitted word for 2017: SemEval-2017 Task4: Sentiment Analysis in Twitter

http://www.aclweb.org/anthology/S17-2088

WEI CHEN

Oct.15th

Dataset	Domain	Approach	Task	Level	Evaluation	Reference
crawled and downloaded the first 100 reviews from Amazon & C|net.com	Product reviews	dictionary-based (P-support pruning)	Polarity	Sentence-based	Accuracy 0.79	https://ocs.aaai.org/Papers/AAAI/2004/AAAI04-119.pdf
Crawled from Epinions.com (not clear)	product reviews	unsupervised method, called Opinion Digger	5-star rating		Ranking Loss: 0.49	https://dl.acm.org/citation.cfm?id=1871739
Reviews crawled tripadvisor.com or zagats.com	Restaurant and hotel reviews	lexicon-based classifiers	Polarity (pos/neg)		precision: 68.0% / 77.2% recall: 90.7% / 86.3%	http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.182.4520&rep=rep1&type=pdf
Chinese restaurant reviews	Restaurant reviews	unsupervised approach to aspect-based opinion polling	ternary (pos/neg/neu) Different level	aspect-sentiment extraction	75.5%	https://dl.acm.org/citation.cfm?id=1646233
Data crawled from the prevalent forum Web sites, including cnet.com, viewpoints.com, reevoo.com and gsmarena.com.	product reviews	SVM sentiment classifier	binary	aspect-sentiment extraction	F1: 71.7%-85.1%	http://anthology.aclweb.org/P/P11/P11-1150.pdf
(MPQA) corpus 535 newswire documents manually annotated	newswire documents	Incorporate structural inference motivated by compositional semantics into the learning procedure.	binary	phrase-level	accuracy: 90.70%	https://www.cs.cornell.edu/home/cardie/papers/emnlp08.pdf

OpenTable.com CitySearch.com TripAdvisor hotel review	Restaurant and hotel reviews	unsupervised and weakly supervised topic modelling approaches: Majority LDA MG-LDA STM Local LDA SVM	5-star rating	multi-aspect sentence level	LAE: 0.560 - 0.790 hotel [40] reviews	http://www.myleott.com/MASA_SENTIRE2011.pdf
Reviews for Mp3 players from Google Product Search4 and subsets of reviews of hotels and restaurants from Google Local Search.	product, hotel, restaurant reviews	Unsupervised. Multi-grain Topic Models: MG-LDA	binary	Aspect based Document level	Ranking Loss: 0.669	https://arxiv.org/pdf/0801.1063.pdf
reviews of five electronics products Amazon.com and C|net.com	product reviews	Unsupervised *Pattern Learning* *Subclass Extraction*	ternary		precision: 84.8% recall: 89.28%	https://ac.els-cdn.com/S0004370205000366/1-s2.0-S0004370205000366-main.pdf?_tid=55aa263e-b83b-11e7-bdba-00000aab0f27&acdnat=1508795243_692d509e9c24baf51271a4cd0e81e60b
reviews from epinions.com on automobiles, banks, movies, and travel destinations	product reviews	Unsupervised PMI-IR algorithm	binary	Document level	Overall Accuracy 74.39 % correlation 0.5174	https://arxiv.org/ftp/cs/papers/0212/0212032.pdf

IMDb reviews http://www.imdb.com/reviews/index.html	Movie reviews	supervised learning Naïve Bayes Maximum entropy Support vector machines	binary	Document level	Best accuracy 83%	https://arxiv.org/pdf/cs/0205070.pdf 2002
English language versions of foreign news documents from FBIS	Subjectivity classification	bootstrapping approach	Binary	Sentence based	Recall 32.9 Precision 91.3	http://ccc.inaoep.mx/~villasen/index_archivos/cursoTATII/ClasificacionOpiniones/Riloff-ExtractionPatternsForSubjetiveExpresions03.pdf
	movie reviews	hierarchical multi-classifier	three and four star rating	document-level multi-way sentiment detection		https://www.researchgate.net/profile/Ingrid_Zukerman2/publication/221102026_A_Hierarchical_Classifier_Applied_to_Multi-way_Sentiment_Detection/links/00b7d51807879f1221000000.pdf
Amazon product review dataset 82 million product reviews	Product reviews				Over 92%	https://arxiv.org/pdf/1704.01444.pdf from Open AI

For supervised learning

Key: feature engineering. A large set of features have been tried by researchers. E.g., Terms frequency and different IR weighting schemes Part of speech (POS) tags Opinion words and phrases Negations Syntactic dependency, domain adaption and cross-lingual