

Detecting General Opinions from Customer Surveys

Evgeny A. Stepanov & Giuseppe Riccardi

Signals & Interactive Systems Lab
University of Trento, Italy
{stepanov,riccardi}@disi.unitn.it

SENTIRE ICDM 2011-12-11

Outline

- 1 Introduction
 - Motivations
- 2 Methodology
- 3 Experiments and Results
 - Telecom Italia Data Set
 - General Opinion Detection
 - Polarity Classification
- 4 Conclusion

Motivations

- **Problem:** Pairing of free-text comments with user ratings
 - User rating might not reflect the true “rating” of the text
 - additional comments
 - comments on only negative aspects
- **Effect:** Inaccurate models for classification
- **Solution:** Identify *general* opinions

Examples

General Negative Opinion (user rating 4):

Il sistema **product_i** va bene per la clientela Consumer e Business di piccole dimensioni. E' del tutto inadeguato per la clientela Corporate.

*The system **product_i** is good for small size Consumer and Business customers. It is totally inadequate for Corporate customers.*

Performance Problem (user rating 3):

Il problema maggiore e' sicuramente la lentezza del sistema in generale nell'effettuare transazioni.

The biggest problem, certainly, is the general slowness of the system in making transactions.

Questions to Answer

- Are user provided ratings useful to identify documents bearing general opinions?
- Are user ratings effective in predicting opinion polarity of their free-text comments?

User Comment Annotation

Manual Annotation Categories (15)

- 11 labels for different product aspects users were reporting problems about;
- 1 label for the suggestions;
- 2 labels for general opinions: positive and negative;
- 1 label for everything else, i.e. others;

Classification Categories (4); Automatic grouping

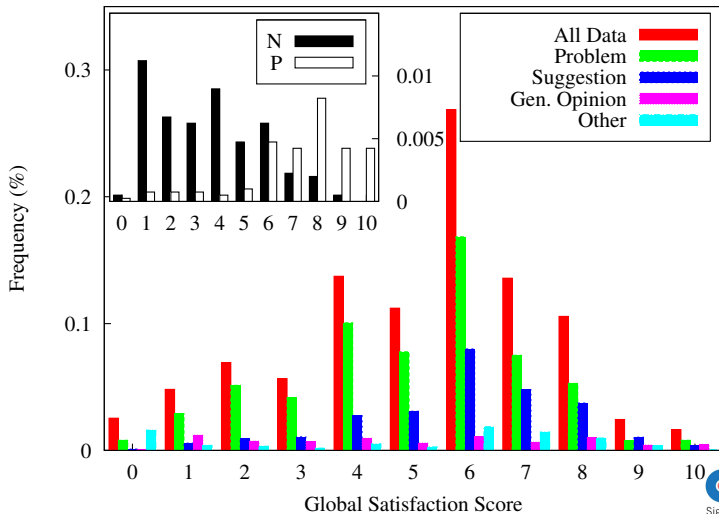
- Problems
- Suggestions
- General Opinions
- Other

Distribution of Categories

Category	Training		Testing	
	Count	Percentage	Count	Percentage
Problems	2324	57.9%	389	67.4%
Suggestions	926	23.1%	67	11.6%
Gen. Opinions	266	6.6%	19	3.3%
Other	321	8.0%	41	7.1%
Problem & Gen. Opinion	40	1.0%	20	3.5%
Suggestion & Gen. Opinion	10	0.3%	1	0.2%
Problem & Suggestion	125	3.1%	39	6.8%
Problem & Other	0	0.0%	1	0.2%
Total	4012	100.0%	577	100.0%



Global Satisfaction Score Distribution



Real Distribution Training: Binary (Av. F1)

	Classifier	CV5	Test Set
GSS	SVM	0.000	0.000
	BoosTexter	0.000	0.000
Unigram	SVM	0.055	0.000
	BoosTexter	0.288	0.224
Unigram + GSS	SVM	0.000	0.000
	BoosTexter	0.307	0.233

- GSS-only – 0 recall
- GSS effect on unigram classifier is not significant

Real Distribution Training: 4-way (Av. F1)

Classifier		CV5	Test Set
GSS	BoosTexter	0.000	0.000
Unigram	BoosTexter	0.320	0.317
Unigram + GSS	BoosTexter	0.341	0.309

- GSS-only – 0 recall
- Positive GSS effect on unigrams is not significant in CV5
- Effect is not seen on a Test Set

Balanced Training: Binary (Av. F1)

	Classifier	CV5	Test Set
GSS	SVM	0.538	0.355
	BoosTexter	0.631	0.194
Unigram	SVM	0.671	0.212
	BoosTexter	0.691	0.168
Unigram + GSS	SVM	0.689	0.187
	BoosTexter	0.701	0.175

- Data Size: $2 * 250$
- Non-zero GSS-only performance – altered distribution

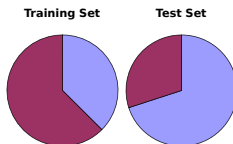
Balanced Training: 4-way (Av. F1)

	Classifier	CV5	Test Set
GSS	SVM-MV	0.444	0.114
	BoosTexter	0.422	0.182
Unigram	SVM-MV	0.796	0.171
	BoosTexter	0.544	0.225
Unigram + GSS	SVM-MV	0.765	0.154
	BoosTexter	0.581	0.235

- Data Size: $4 * 250$
- SVM-MV: Majority Voting
- BoosTexter vs. SVM-MV: different GSS distributions

General Opinion-Only Data Set

Polarity	Training		Test	
Positive	118	37.3%	28	70.0%
Negative	197	62.7%	12	30.0%
Mixed	1	0.3%	0	0.0%
Total	316	100.0%	40	100.0%



Real Distribution Training (Av. Accuracy)

	Classifier	CV5	Test Set
GSS	SVM	0.835	0.805
	BoosTexter	0.813	0.745
Unigram	SVM	0.703	0.525
	BoosTexter	0.758	0.595
Unigram + GSS	SVM	0.842	0.795
	BoosTexter	0.819	0.745

- CV5: GSS > Unigram
- CV5: GSS + Unigram > GSS; not significant
- Data Size is small!

Balanced Distr. Training: 1-10 (Av. Acc.)

	Classifier	CV5	Test Set
GSS	SVM	0.791	0.805
	BoosTexter	0.778	0.780
Unigram	SVM	0.644	0.565
	BoosTexter	0.713	0.590
Unigram + GSS	SVM	0.804	0.820
	BoosTexter	0.804	0.775

- Data Size: $2 * 115$

Polarity Classification: Summary of the Results

- GSS is very effective
 - GSS-only: removing the “ambiguous” middle part has a positive effect
 - Random Under-sampling has a negative effect on unigram-only and CV5
 - Unigram + GSS > GSS on the balanced data with “ambiguous” middle range documents
- GSS only classifier trained on balanced data with removed “ambiguous” range can predict document polarity sufficiently well.
 - Unsupervised polarity classification with thresholding is reasonable for general opinions.

