A Multilingual Polarity Classification Method Using Multi-Label Classification Technique Based on Corpus Analysis

Yohei Seki
Toyohashi University of Technology (Visiting in Columbia University)
seki@ics.tut.ac.jp

Research Goal: Clarify Effective Polarity Classification Clues and Methods

In *NTCIR-7 MOAT*, we have new challengs as follows:
・Participants could use *NTCIR-6 OAT* corpus for training.
・Many participants focused on language portable approaches.

Based on these, our research goal in NTCIR-7 is as follows:
・Using the features that were acquired from the significance test in *NTCIR-6 OAT* and *MPQA* corpora, we estimated the effectiveness in opinion detection and polarity classification.
・In NTCIR-7 *MOAT*, polarity classification was the problem to classify three lables: positive, negative, or neutral. We compared two multi-label classification techniques:
A.) *SVM voting* and B.) *Mulan (label power set classification)* .

System Overview

・My system overview is shown in Figure 1. The common architecture was implemented both in Japanese and English.
・The polarity classification, opinion holder identification, and relevance judgment module were based on the results from opinion detection module.
・The opinion detection system was based on the differentiation of author opinions and authority opinions [1].

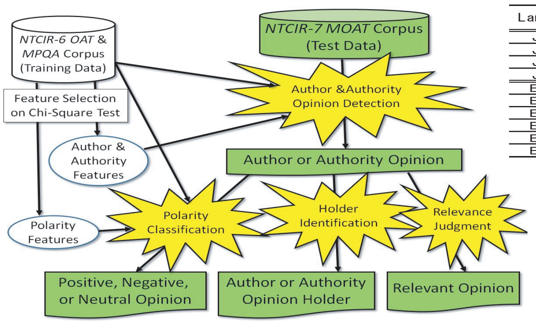

Figure 1. System Flow Overview

Feature Selection with Corpus Analysis

・The features in opinion detection and polarity classification both in Japanese and English were selected based on $\chi 2$ test (the significance probability was 5%, two-sided test) on *NTCIR-6 OAT* and *MPQA* corpora.
In polarity classification case, if a feature appeared more frequently in the sentence with one polarity type than sentences with other two polarity types, it was selected.
To avoid the error from low frequency data, we only investigated the features which appeared more than five times in the *NTCIR-6 OAT* corpus.
・As feature types, we focused on a.) terms or term types through the abstraction using thesaurus (*Bunrui-Goi Hyo* or *WordNet*) or lexicon (*General Inquirer*, *Wiebe*'s subjective lexicon, *Hatsivassiloglou*'s adjective entries etc.) and b.) syntactic pairs using dependency parsers (*Minipar*, *Cabocha*).
・In Japanese, we utilized grammatical subjects, action semantic primitives, syntactic pairs, and keywords as feature types.
・In English, we utilized subjective term lexicons and two syntactic pairs such as "subject"-"verb" and "auxiliary verb"-"verb" relationship using *Minipar* parser.
・The selected features are shown in Tables 3-6 in the paper.

Polarity Classification Using Multi-label Classification Techniques

We implemented two multi-label classification techniques.
1. The first one was a voting approach with three *SVM* classifiers. The features selected were used for each classifier. This approach was implemented using *SVMlight*.
2. Another one was a multi-label classifier using *Mulan* (label power set classifier) system. Note that we could not differentiate the feature sets according to three polarity types in this classifier, so we combined them into one feature set.

Experimental Results in NTCIR-7

・For polarity classification in Table 1, the results using *SVM voting* approach were shown as RunID 1 and the result using *Mulan* classifier was shown as RunID 2 in Japanese and as RunID 3 in English.
・The results of *SVM voting* approach were better than the results of *Mulan*. Note that *SVM approach* need to tune cost parameters according to each classifier and we tuned them by using sample data in *NTCIR-7 MOAT*, but we did not tune any parameters in Mulan.

Table 1. Evaluation Results in NTCIR-7 MOAT

Lang	Run ID	L / S	Opinionated			Relevance			Polarity			Opinion Holder		
			P	R	F	P	R	F	P	R	F	P	R	F
J	1	L	0.6742	0.5620	0.6130	0.5527	0.2925	0.3825	0.4596	0.2140	0.2920			
J	2	L	--- (same in TUT-1)			--- (same in TUT-1)			0.4283	0.1994	0.2721			
J	1	S	0.5416	0.6199	0.5781	0.3062	0.3357	0.3203	0.4806	0.2417	0.3216			
J	2	S	--- (same in TUT-1)			--- (same in TUT-1)			0.4535	0.2281	0.3035			
E	1	L	0.3185	0.4092	0.3582	0.2092	0.1755	0.1909	0.1943	0.1830	0.1885	0.3923	0.2833	0.3290
E	2	L	0.3282	0.2562	0.2878	0.1647	0.1136	0.1344	0.1896	0.1142	0.1425	(0.3656)	(0.1689)	(0.2311)
E	3	L	--- (same in TUT-1)			--- (same in TUT-1)			0.1621	0.1527	0.1573	--- (same in TUT-1)		
E	1	S	0.0961	0.4149	0.1561	0.0740	0.1853	0.1057	0.0569	0.2180	0.0903	0.1250	0.2829	0.1735
E	2	S	0.1039	0.2724	0.1504	0.0615	0.1220	0.0817	0.0484	0.1185	0.0687	(0.1257)	(0.1821)	(0.1487)
E	3	S	--- (same in TUT-1)			--- (same in TUT-1)			0.0359	0.1374	0.0569	--- (same in TUT-1)		

Discussion: Improve Multilabel classification?

・We concluded the reason why *SVM* results were over than *Mulan* results was that we could not discriminate the feature sets each by polarity type in *Mulan*.
・We also investigated a confusion matrix from *SVM* voting and *Mulan* as in Table 2.
・You could confirm that the results using *Mulan* classifier were sometimes better than the results using *SVM* classifier, for example, negative classifier in Japanese.
・In future, we plan to implement Multi-label classification technique to discriminate three polarity types as inputs.

Table 2. Confusion matrix with SVM voting and Mulan approaches

Lang	Method			Assessment (Lenient)		
				Pos	Neg	Neu
J	SVM voting	System	Pos	15	3	51
			Neg	9	66	349
			Neu	18	52	329
			(No)	63	173	788
	Mulan		Pos	15	12	105
			Neg	16	89	346
			Neu	11	20	278
			(No)	63	173	788
E	SVM voting		Pos	18	30	4
			Neg	64	136	18
			Neu	25	37	3
			(No)	165	318	40
	Mulan		Pos	18	17	2
			Neg	49	102	12
			Neu	40	84	11
			(No)	165	318	40

Reference

[1] Y. Seki, N. Kando, and M. Aono: Multilingual Opinion Holder Identification Using Author and Authority Viewpoints, Information Processing & Management(to appear)

The Seventh NTCIR Workshop Meeting (NTCIR-7)
December 16-19, 2008. NII, Tokyo, Japan