

Topic Classification of Spoken Inquiries Using Transductive Support Vector Machine

Rafael Torres[†], Hiromichi Kawanami[†], Tomoko Matsui[‡], Hiroshi Saruwatari[†], Kiyohiro Shikano[†]

[†] Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan

[‡] Department of Statistical Modeling, The Institute of Statistical Mathematics, Tokyo, Japan

Abstract: We investigate the topic classification of spoken inquiries received by a guidance system with a semi-supervised learning approach based on TSVM. We obtained gains in classification performance for specific topics by using this approach, in comparison to those of its supervised counterpart.

1. Research Background

- Supervised learning for topic classification requires manually labeled data.
- Manual data labeling is a costly process.
- Unlabeled data are usually abundant and cheap to obtain.
- Using unlabeled data to improve topic classification performance is desirable.

- Semi-supervised learning allows to use labeled and unlabeled data.

2. Takemaru-kun System

- A speech-oriented guidance system placed at Ikoma-City North Community Center
- Operating daily since November, 2002
- One-question one-response dialog strategy

Offers guidance about:

- Center facilities and services
- Local sightseeing
- Weather forecast, time, news
- Others



3. Transductive Support Vector Machine (TSVM)

Training vectors:

$$\mathbf{x}_i \in \mathbb{R}^n, i = 1, \dots, l$$

$$\mathbf{x}_j^* \in \mathbb{R}^n, j = 1, \dots, m$$

Classes: $y_i \in \{1, -1\}$

$y_j^* \in \{1, -1\}$

Algorithm

- Begins with labeling the unlabeled data based on the prediction of a regular (inductive) SVM.
- Then it improves the solution by switching the labels of these samples so that the objective function decreases.

Input: Labeled training samples $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)$

Parameters: C, C^*

Unlabeled training samples $\mathbf{x}_1^*, \dots, \mathbf{x}_m^*$

num_+ : number of unlabeled samples for class +

- The influence of unlabeled samples is increased by incrementing the cost factors C_+, C_-^* , initialized with a very small value, up to the defined value C^* (loop 1)
- Switch class labels of two samples if it leads to a decrease in the objective function (loop 2)

```
while (( $C_+^* < C^*$ ) || ( $C_-^* < C^*$ )) {
    ( $\mathbf{w}, b, \xi, \xi^*$ ) = solve_svm([(x1, y1), ..., (xl, yl)], [(x1*, y1*), ..., (xm*, ym*)], C, C+, C-*)
    while ((exists q, r: (yq* * yr* < 0) & (xiq* > 0) & (xir* > 0) & (xiq* + xir* > 2))) {
        yq* = -yq*
        yr* = -yr*
        ( $\mathbf{w}, b, \xi, \xi^*$ ) = solve_svm([(x1, y1), ..., (xl, yl)], [(x1*, y1*), ..., (xm*, ym*)], C, C+, C-*)
        C+* = min(C+* * 2, C*)
        C-* = min(C-* * 2, C*)
    }
}
```

4. Experiments

- Topic classification of ASR results of spoken inquiries from children and adults.

Table 1: Characteristics of the labeled datasets

Topic	Utterance Example	Amount of Samples				ASR Word Acc. (%)	
		Children Training	Children Test	Adults Training	Adults Test	Children Training	Adults Training
chat-compliment	You're so cool.	2548	1066	766	194	63.74	78.11
info-service	Which are the library service hours?	884	206	494	89	73.57	91.36
info-news	Show me today's news.	529	144	484	137	89.95	96.19
info-local	Where is NAIST?	709	187	553	70	74.32	87.47
info-facility	Where is the Habataki Hall?	5007	1653	1795	299	77.39	93.24
info-city	Where is Ikoma's City Hall?	1006	317	504	93	77.13	92.78
info-weather	Which is today's weather forecast?	2947	1073	1099	257	81.13	93.84
info-time	What time is it?	3911	898	984	187	56.66	56.23
info-sightseeing	Where is Kurondo Pond?	647	142	668	79	80.18	90.65
info-access	Where is the closest train station?	681	142	676	83	69.95	89.39
greeting-end	Good bye.	4535	2125	912	269	49.61	73.26
greeting-start	Good morning.	6845	2629	2672	723	54.61	84.46
agent-name	What's your name?	5381	1574	1309	254	69.05	87.49
agent-likings	Which is your favorite food?	4418	2260	851	194	83.88	93.25
agent-age	How old are you?	3446	1108	664	157	78.05	93.90
		43494	15524	14431	3085	72.95	88.42

Table 2: Characteristics of the unlabeled datasets

Unlabeled Dataset	Amount of Samples	
	Children Training	Adults Training
#1 (2005.04 to 2005.12)	119322	110537
#2 (2005.04 to 2006.12)	271744	252428
#3 (2005.04 to 2007.12)	413144	385165

Table 3: Experimental Conditions

Features	Character 1+2+3-grams
SVM, TSVM tool	SVM ^{light} 6.02
Kernel function	RBF
Multi-class classification	one-vs-rest

Results

Table 4: Averaged F-measure results (%) (open test)

Training Dataset Combination	Children	Adults
Labeled only (SVM)	83.54	93.03
Labeled dataset + Unlabeled dataset #1 (TSVM)	83.02	91.75
Labeled dataset + Unlabeled dataset #2 (TSVM)	84.17	92.86
Labeled dataset + Unlabeled dataset #3 (TSVM)	84.28	92.81

Table 5: F-measure results per topic (%) (open test)

Topic	Children SVM	Children TSVM	Adults SVM	Adults TSVM
chat-compliment	64.24	66.91	86.35	81.64
info-service	58.06	59.04	87.65	87.12
info-news	88.89	92.47	95.52	96.30
info-local	56.71	59.50	83.08	84.44
info-facility	82.70	82.15	89.36	89.36
info-city	67.06	73.37	84.34	88.27
info-weather	83.89	85.40	95.46	95.74
info-time	89.67	90.83	95.56	96.83
info-sightseeing	74.74	75.34	91.39	92.21
info-access	44.58	44.89	84.39	92.02
greeting-end	84.87	83.81	93.48	92.17
greeting-start	91.64	91.88	97.76	97.83
agent-name	79.15	80.84	92.15	89.50
agent-likings	89.75	90.62	93.30	91.64
agent-age	89.26	89.69	95.71	95.68
Averaged	83.54	84.28	93.03	92.81

Conclusions

- Most of the topics presented classification performance improvements for children, and more than half of the topics in the case of adults.
- In particular, we found improvements in classification performance for topics such as city information with 6.30% of F-measure improvement for children, and access information with 7.63% for adults.