

음성인식을 이용한 자동 호 분류 철도 예약 시스템

심유진(KT), 김재인(KT), 구명완(KT)

<차 례>

- | | |
|----------------|---------------|
| 1. 서론 | 3. 호 분류 실험 |
| 2. 호 분류 알고리즘 | 3.1. 데이터베이스 |
| 2.1. 입력 벡터 추출 | 3.2. 실험환경 |
| 2.2. 분류 행렬 생성 | 4. 실험 결과 |
| 2.3. 유사도 생성 과정 | 5. 결론 및 개선 방향 |

<Abstract>

A Train Ticket Reservation Aid System Using Automated Call Routing Technology Based on Speech Recognition

Yu-Jin Shim, Jae-In Kim, Myung-Wan Koo

This paper describes the automated call routing for train ticket reservation aid system based on speech recognition. We focus on the task of automatically routing telephone calls based on user's fluently spoken response instead of touch tone menus in an interactive voice response system. Vector-based call routing algorithm is investigated and mapping table for key term is suggested. Korail database collected by KT is used for call routing experiment.

We evaluate call-classification experiments for transcribed text from Korail database. In case of small training data, an average call routing error reduction rate of 14% is observed when mapping table is used.

* Keywords : Speech recognition, Automatic call routing, Train reservation system

1. 서 론

음성인식이 보편화되면서 사람들은 기존에 버튼이나 문자를 이용하여 해야 되는 일을 음성으로 대체하기 시작했다. 음성인식 서비스를 이용함에 있어서 사람들은 기계가 사람이 자연스럽게 말하는 문장까지도 잘 알아듣기를 바란다. 그런 맥락에서 사람들은 자유롭게 발화된 문장을 가지고 그 말의 의도를 파악하고 반응할 수 있는 기계를 만드는데 관심을 가지게 되었다.

그것의 응용분야 중 하나가 호 분류로 볼 수 있는데 사용자가 전화를 걸어 자신의 의도를 말하면 기계는 그 사람이 원하는 것이 무엇인지를 파악하여 그런 일을 처리할 수 있는 부서로 전화 호를 전달해 주는 것이다. 이러한 서비스는 주로 콜센터나 안내시스템과 같이 현재는 안내원들이 직접 그 역할을 수행하고 있으나 인건비 문제나 사용자의 편리성 관점에서 사람의 역할을 대신하여 편리하게 사용될 수 있다.

미국의 통신회사인 AT&T에서는 “How May I Help You?” 라는 서비스를 제공하는데 이는 자사의 고객센터에 걸려오는 전화를 안내원을 대신하여 기계가 처리해 주는 서비스로써 인건비 절감의 차원에서 좋은 결과를 가져왔다고 알려져 있다[3].

이와 같이 호 분류를 구현함에 있어서 보다 좋은 성능을 내기 위하여 많은 연구가 이루어지고 있다. 분류 정확도를 높이기 위해 여러 가지 분류기(classifier)를 조합하여 은행의 콜 분류 시스템에 적용하여 그 성능을 보기도 하고[6] 요소 단어나 어절 대신 음소 열을 기본단위로 하고 각각의 카테고리 별로 다른 언어모델과 분류기를 사용하는 방법도 제안되고 있다[7]. 또한 일반적으로 분류기는 모든 클래스의 중요도를 동일하게 보는데 실제 호 분류에 있어서는 분명 더 중요한 카테고리가 존재한다는 점에서 훈련 시에 중요한 카테고리에 비중 실어주는 방법도 연구되고 있다[5].

본 논문에서는 자동 호 분류 알고리즘 중에서도 벡터 기반 방법에 기반을 두었고 요소어절을 사용함에 있어서 비슷한 의미를 가지는 것은 하나의 그룹으로 묶을 수 있는 테이블을 사용하여 그 방법이 성능을 향상시킬 수 있는지 철도 예약 시스템을 대상으로 실험해 보았다.

본 논문의 구성은 다음과 같다. 1장의 서론에 이어 2장에서는 호 분류를 구현함에 있어서 입력 벡터를 추출하고 분류 행렬을 생성하는 방법 및 유사도를 계산하는 방법에 대해 설명하였다. 3장에서는 본 연구에서의 분류 실험에 쓰인 데이터 베이스와 실험과정에 대해 기술한다. 4장에서는 분류 실험 결과를 제시한 후 5장에서 결론을 맺는다.

2. 호 분류 알고리즘

호 분류의 과정은 크게 세단계로 나눌 수 있다. 첫 번째는 호 분류에 영향을 미칠 수 있는 단어의 리스트를 정하고 그것을 이용하여 입력 벡터를 추출하는 것이고 두 번째는 훈련 시에 분류 매트릭스를 생성하는 하는 것, 세 번째는 카테고리를 정하는 데 있어서 유사도(similarity score)를 생성한다.

2.1. 입력 벡터(Query Vector)추출

입력 벡터는 입력을 해당하는 카테고리로 분류되면서 다른 카테고리에 해당하는 입력은 거절해야 하므로 분별력을 가지는 것이 중요하다. 보통 입력으로 쓰이는 문장은 음성인식의 결과 이므로 사람이 발화한 그대로라고 보장 할 수 없기에 여러 환경에서 강인해야 한다. 따라서 입력된 모든 단어를 입력 벡터의 요소로 쓰지 않고 분류에 영향을 미치는 단어만을 입력 벡터의 요소로 쓰는 것이 필요하다. 본 논문에서는 입력된 문장을 어절 단위로 정렬하여 “예, 저, 해주세요” 등과 같이 분류에 도움이 되지 않는 단어를 삭제하는 방법으로 입력 벡터 요소 어절(relevant term)로 분류하였다. 또한 비슷한 단어로 구성되어 형태소적 의미가 같은 경우는 같은 요소 어절에 해당하도록 mapping table을 만들어 사용하였다. 이러한 과정을 통하여 훈련이 용이하고 또한 제한된 훈련데이터를 효율적으로 활용할 수 있게 된다. <표 1>은 mapping table의 한 예이다. 또한 비슷한 단어로 구성되어 형태소적 의미가 같은 경우는 같은 요소 어절에 해당하도록 mapping table을 만들어 사용하였다. 이러한 과정을 통하여 훈련이 용이하고 또한 제한된 훈련데이터를 효율적으로 활용할 수 있게 된다. <표 1>은 mapping table의 한 예이다.

<표 1> mapping table의 예

어절	해당 요소 어절
예약	1
예약을	1
얼마예요	2
얼마죠	2
...	...
끊으려고	3

이러한 테이블을 통과한 요소 어절의 개수가 호 분류 시스템에 있어서 사용되

는 어휘의 수가되어 입력 벡터(Query vector)의 차수가 된다.

그러므로 입력된 텍스트에 요소 어절이 한번 나타나면 그 값이 1이 되고 두 번 나타나면 2 그 외 나타나지 않는 요소의 값들은 0으로 구성된 입력 벡터가 생성되어 훈련과 테스트에 있어서 입력으로 사용되게 된다.

2.2. 분류 행렬(Routing Matrix)생성

분류를 훈련시키는 방법에는 여러 가지가 있으나 본 논문에서는 벡터 기반 방법을 사용하였다[2].

일단 요소 어절을 이용하여 입력 벡터가 생성되면 그와 함께 그에 해당하는 카테고리 정보를 가지게 된다.

이때 n 번째 카테고리에 해당하는 입력 벡터들을 모아서 모두 더하여 축척 빈도 벡터(accumulated count vector, A_n)를 생성한다. 만약 n 개의 카테고리가 있고 입력 벡터의 차수 즉 요소어절의 개수가 m 개라면 $m \times n$ 행렬인 $A_{t,d}$ 를 생성할 수 있다.

또한 요소 어절의 빈도수를 그대로 쓰게 되면 훈련 데이터에 많이 나오는 어절에 비중이 주어지게 되므로 이를 보상하기 위하여 각각의 카테고리에서 나타난 빈도수로 아래의 식 (1)과 같이 정상화(normalization)하는 과정을 거친다.

$$B_{t,d} = \frac{A_{t,d}}{\left(\sum_{1 \leq e \leq n} A_{t,e}^2 \right)^{\frac{1}{2}}} \quad (1)$$

또한 여러 카테고리에 해당하는 요소 어절 보다는 하나의 카테고리에만 해당하는 요소 어절이 호 분류에는 변별적인 측면에서 중요하므로 요소어절이 발생하는 카테고리의 수로 각 요소어절에 해당되는 벡터 값을 아래의 식 (2)과 같이 나눠준다[1].

$$IDF(t) = \log_2 \frac{n}{d(t)} \quad (2)$$

이때 t 는 term으로 요소 어절을 뜻하고 n 은 훈련데이터에서 나타난 입력 문장의 수, $d(t)$ 는 요소어절 t 를 포함하는 입력 문장의 수이다. $B_{t,d}$ 의 각 행에 $IDF(t)$ 를 곱하면 아래의 식 (3)과 같이 행렬 C 를 생성된다.

$$C_{t,d} = IDF(t) \cdot B_{t,d} \quad (3)$$

2.3. 유사도(similarity) 생성 과정

이제 입력된 벡터가 어떤 카테고리에 해당하는지 판단하는 과정이 필요하다.

요소 어절 벡터와 카테고리 벡터에 대해 균일한 표현을 함과 동시에 차수를 줄이기 위해 singular vector decomposition이 적용되었다. 이때 요소 어절의 개수를 m 이라 하고 카테고리의 개수를 n 이라고 하면 아래의 식 (4)과 같이 표현될 수 있다.

$$C = U \cdot S \cdot V^T \quad (4)$$

r 이 행렬 C 의 랭크 값일 때 U 는 $m \times r$ orthonormal 행렬이 되고 V 는 $r \times n$ orthonormal 행렬, S 는 양수로만 이루어진 $r \times r$ 행렬로 0아닌 값인 $s_{1,1}, \dots, s_{r,r}$ 이 아래 차순으로 정렬되어 있는 diagonal 행렬이다.

행렬 U 의 i 번째 행은 i 번째 요소어절을 표현하는 r 차 벡터가 되고 행렬 V 의 j 번째 행은 j 번째 카테고리를 표현하는 r 차의 벡터가 된다.

따라서 $m \times 1$ 의 입력 벡터를 Q 라고 하면 아래의 식 (5)과 같이 pseudo-document 행렬 D 를 생성할 수 있다[4].

$$D = U^T Q \quad (5)$$

유사도는 위에서 구해진 pseudo-document 행렬 D 와 scaled document 행렬인 SV^T 사이의 cosine score로 구해질 수 있으며 아래식 (6)과 같다.

$$\cos(X, Y) = \frac{X \cdot Y^T}{\sqrt{\sum_{1 \leq i \leq n} x_i^2 \cdot \sum_{1 \leq i \leq n} y_i^2}} \quad (6)$$

식(6)에 의한 유사도의 결과는 n 차의 벡터가 될 것이고 j 번째의 값은 입력 벡터와 j 번째 카테고리의 유사도를 나타내게 된다.

따라서 가장 큰 값을 가진 카테고리가 주어진 입력에 해당하는 카테고리로 추정되게 된다.

3. 호 분류 실험

3.1. 데이터베이스

실험을 위한 데이터베이스로는 현재 상용화된 음성인식 철도 예약 시스템의

구축을 위해 KT측에서 수집한 Korail DB를 사용하였다. 이 데이터베이스는 철도 예약을 음성으로 하면서 사용자의 질의 및 동의로 이루어져 있는데 본 실험에는 사용자가 처음 전화를 걸었을 때 본인의 의도를 말하는 한 문장 만 입력으로 사용되었다. 이러한 문장 셋을 데이터로 사용되는데 각각의 문장에는 사용자가 전화를 건 목적이 철도 예약을 위한 것인지, 요금이나 시간들을 확인하기 위한 것인지, 예매한 표를 취소하기 위해서인지, 단순 안내를 바라는 것인지에 대한 분류 카테고리 정보도 덧붙여진다. 훈련 데이터는 200개로써 100개로 훈련했을 경우와 200개로 훈련했을 경우 성능의 변화를 살펴보고 테스트 데이터는 전체 200개인데 이것을 4개의 셋으로 나누어 각각 테스트 해 보았다.

3.2. 실험 환경

먼저 훈련에 사용되는 입력 리스트가 정해지면 어절 분리 프로그램을 이용하여 중복되지 않는 어절의 집합을 찾아내게 된다. 그러면 그 어절의 집합 중에서 호 분류와 상관없는 단어는 삭제하고 새로운 어절리스트를 생성한 것이 2.1절에서 언급되었던 입력 벡터를 추출하는데 사용되는 요소 어절이 된다.

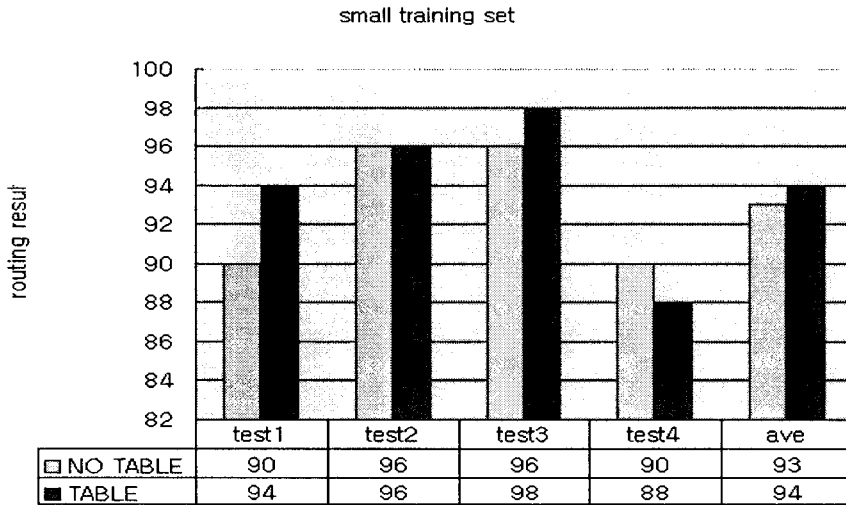
이 과정에서 형태소적 의미가 같은 경우는 같은 요소 어절에 해당하도록 실험을 할 수 있는 mapping table을 만들게 된다. 그러면 위에서 생성한 요소어절 리스트와 mapping table을 이용하여 축적 빈도 벡터를 만들어 정상화하는 과정을 거친 다음 singular vector decomposition을 하고 유사도를 계산하여 훈련이 잘 되었는지 확인하게 된다.

테스트 할 때에는 훈련에서 얻어진 행렬을 이용하여 유사 도를 계산하여 주어진 입력이 어떤 카테고리에 해당하는지를 추정하게 된다.

4. 호 분류 실험 결과

먼저 훈련 데이터를 100개로 맵핑 테이블을 쓰는 경우와 아닌 경우를 비교해서 실험을 해 보았다.

훈련 데이터에 대해 분류 성능을 보면 맵핑 테이블을 쓰지 않고 요소어절을 그대로 사용할 경우는 100%인데 비해 맵핑 테이블을 쓸 경우는 99%로서 다소 떨어졌다. 그러나 테스트 시에는 아래 <그림 1>에서 보이는 바와 같이 테스트 셋에 따라 경향의 차이는 있지만 맵핑 테이블을 쓰는 경우가 더욱 좋은 성능을 내어서 테스트 결과 평균 오분류 감소도(ERR)의 값이 14%가 된다는 것을 관찰 할 수 있었다.

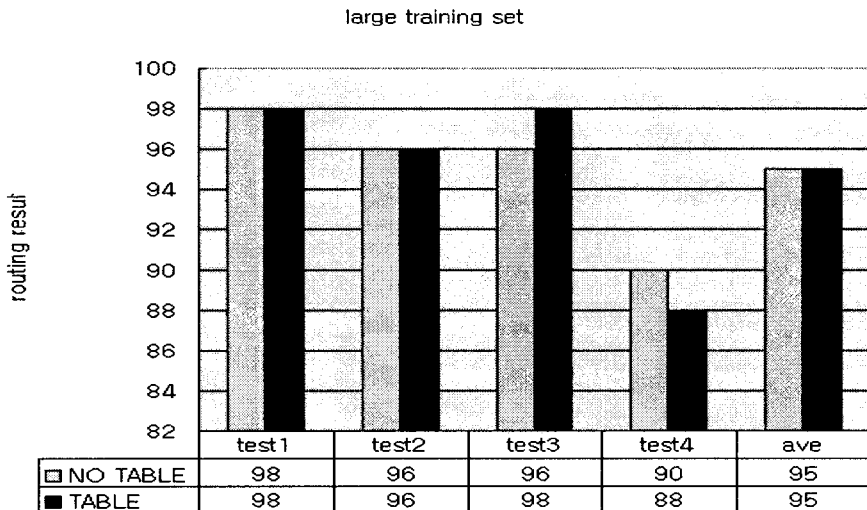


<그림 1> 훈련 데이터량이 적은 경우 분류율

그러나 훈련 데이터를 200개로 해서 같은 실험을 할 경우 훈련 데이터에 대한 분류 결과는 작은 훈련 데이터로 실험 했을 때와 같이 테이블 없이 요소어절만 사용한 경우 그 결과가 더 좋았으나 테스트 결과는 아래 <그림 2>와 같이 맵핑테이블을 사용한 경우와 사용하지 않을 경우나 같게 나왔다. 실험결과로 볼 때 맵핑테이블을 사용하는 방법은 훈련 데이터의 양이 충분하지 않을 경우 좀 더 좋은 성능을 낸다고 볼 수 있다.

또 분류 정확도 면에서 보면 역시 일반적으로 알려져 있듯이 훈련 데이터의 양이 많을수록 테스트 결과가 좋다는 것을 보여준다.

잘못된 분류가 이루어진 데이터를 살펴보면 대부분은 훈련에 나타나지 않은 단어가 테스트 데이터에 있는 경우이고 간혹 한 문장에 나타나는 요소어절이 여러 개인 경우 혼돈되어 다른 카테고리로 분류되는 경우도 있었다. 본 실험의 경우 데이터가 철도 예약 시스템으로써 사용자가 전화를 건 목적이 몇 가지로만 한정할 수 있고 또한 쓰여지는 단어도 광범위 하지 않기에 이 정도의 데이터 양으로도 어느 정도의 성능을 낼 수 있지만 복잡도가 있는 시스템에 호 분류를 적용하기 위해서는 훨씬 많은 데이터가 제공되어야만 원하는 성능을 기대할 수 있을 것이다.



<그림 2> 훈련 데이터량이 많을 경우 분류율

5. 결론 및 개선 방향

본 논문에서는 음성인식을 이용한 자동 호 분류 철도 예약 시스템을 구현하기 위해 필요한 호 분류 실험을 해보았다. 벡터 기반 알고리즘을 사용하여 입력된 문장으로부터 입력 벡터를 추출하고 분류행렬을 생성하여 유사도를 계산하는 과정을 거친 실험을 하였고 입력 벡터를 추출하는 과정에서 맵핑 테이블의 개념을 활용하여 성능의 개선이 이루어 질 수 있는지 살펴보았다. 그 결과 훈련 데이터의 양이 적은 경우 오 분류 감소율이 14% 정도로 효과가 있는 것으로 관찰되었다. 그러나 훈련 데이터가 클 경우는 전반적으로 오 분류율이 감소하면서 맵핑테이블을 사용하는 것과 사용하지 않는 것에 차이가 없다는 것이 관찰되었다.

참 고 문 헌

- [1] Chu-Carroll, J. and Carpenter, R. L. "Vector-based Natural Language Call Routing", *Computational Linguistics*, vol.25, no.3, pp.361-388, 1999.
- [2] Lee, C-H., Carpenter R. L., et al., "A Study on natural language call routing", in *4th IEEE Workshop on Interactive Voice Technology for Telecommunication Applications*, 1998.
- [3] Allen L. Gorin, Alicia Abella, et al., "Automated Natural Spoken Dialog", *IEEE Computer Magazine*, vol. 35 (4), pp. 51-56, April 2002.
- [4] S. Deerwester, S. Dumais, G. Furnas, T. Lamdauer and R. Harshman, "Indexing by Latent

- Semantic Analysis”, *Jour. Amer. Soc. of Information Science*, Vol. 41, pp.391-407, 1990.
- [5] Gokhan Tur, “Cost-Sensitive Call Classification,” Proc. ICSLP, Jeju, Korea, Oct. 2004.
- [6] Imed Zitouni, Minkyu Lee, Hui Jiang, “Constrained Minimization Technique for Topic Identification Using Discriminative Training and Support Vector Machines”, *Proc. ICSLP*, Jeju, Korea, Oct. 2004.
- [7] Qiang Huang, Stephen Cox, “Mixture Language Models for Call Routing”, *Proc. ICSLP*, Jeju, Korea, Oct. 2004.

접수일자: 2004년 11월 15일

게재결정: 2004년 12월 10일

▶ 심유진(Yu-Jin Shim)

주소: 137-792 서울특별시 서초구 우면동 17

소속: KT 서비스개발연구소

전화: 02) 526-6762

E-mail: ilovelatte@kt.co.kr

▶ 김재인(Jae-In Kim)

주소: 137-792 서울특별시 서초구 우면동 17

소속: KT 서비스개발연구소

전화: 02) 526-5093

E-mail: jaeinkim@kt.co.kr

▶ 구명완(Myoung-Wan Koo)

주소: 137-792 서울특별시 서초구 우면동 17

소속: KT 서비스개발연구소

전화: 02) 526-5090

E-mail: mwkoo@kt.co.kr