

(KCC2015 우수논문)
 분류 우선순위 적용과 후보정 규칙을 이용한
 효과적인 한국어 화행 분류[†]
 (Effective Korean Speech-act Classification Using
 the Classification Priority Application and a Post-correction Rules)

송 남 훈 [‡] 배 경 만 [§] 고 영 중 [¶]
 (Namhoon Song) (Kyoungman Bae) (Youngjoong Ko)

요약 화행이란 발화 속에 포함되어 있는 화자에 의해 의도된 언어적 행위이다. 대화 시스템에서 입력된 발화에 적합한 화행을 분류하는 것은 중요하다. 기존의 화행분류에 관한 연구는 규칙기반과 기계학습 기반의 방법을 많이 사용한다. 본 논문에서는 대표적인 기계학습 방법인 지지벡터기계(SVM)와 변환기반 학습(TBL)을 조합한 화행 분류 방법을 제안한다. 이를 위해, 화행별 학습 발화의 수에 기반하여 분류 우선순위를 조정함으로써 지지벡터기계의 분류 편향 문제를 해결하였고, 오답일 확률이 높은 분류 결과에 대해서 변환기반 학습을 통해 생성된 보정 규칙을 적용함으로써 화행분류 성능을 개선하는 방법을 제안한다. 본 논문에서 화행별 학습 발화 수의 차이를 고려한 분류 우선순위 변화와 후보정 규칙을 이용한 화행분류 방법을 실험을 통해 평가하였으며, 이는 학습 발화 수가 낮은 화행의 우선순위를 고려하지 않은 기존의 화행 분류보다 성능이 향상되었다.

키워드 : 하이브리드, 화행분류, 지지벡터기계, 변환기반 학습

Abstract A speech-act is a behavior intended by users in an utterance. Speech-act classification is important in a dialogue system. The machine learning and rule-based methods have mainly been used for speech-act classification. In this paper, we propose a speech-act classification method based on the combination of support vector machine (SVM) and transformation-based learning (TBL). The user's utterance is first classified by SVM that is preferentially applied to categories with a low utterance rate in training data. Next, when an utterance has negative scores throughout the whole of the categories, the utterance is applied to the correction phase by rules. The results from our method were higher performance over the baseline system long with error-reduction.

Key words : Ensemble, Speech-Act classifier, Support Vector Machine, Transformation-based Learning

1. 서론

화행이란 발화 속에 포함되어 있는 화자에 의해 의도된 언어적 행위로, 자연어 대화를 처리하는 많은 시스템에서 화자의 의도 파악을 위해 중요한 역할을 한다[1]. 화행은 발화가 수행해야 할 대화 내에서 역할을 결정하기 때문에 대화에 관한 연구에서 화행을 효과적으로 분류하

는 것은 매우 중요하다. 화행 분류를 위해서 기계학습(Machine Learning)을 기반한 다양한 화행 분류 방법들이 연구 되어왔다[2]. 많은 선행 연구들은 지지벡터기계(Support Vector Machine, SVM)를 분류 모델로 사용하고 있으며, 지지벡터기계를 효과적으로 학습하기 위한 학습 자질(feature)에 대한 연구들[3,4,5,6]과 학습 자질을 위한 가중치를 효과적으로 계산하는 연구[7]가 진행되고 있다.

지지벡터기계를 기반으로 하는 분류 모델은 각 화행의 분류기를 학습하기 위해 현재 화행에서 나온 발화들은 최적분류초평면(Optimal Separating Hyperplane, OHS)을 나타내는 방정식을 통해 양(positive)의 분류 점수가 나오고, 현재 화행을 제외한 다른 모든 화행에서 나온 발화들은 음(negative)의 분류 점수가 나오도록 학습을 진행하는 이진분류기이다. 이러한 이진분류기는 그림 1과 같이 하나의 발화가 입력되면 모든 분류기의 분류 점수 중 가장 큰 분류 점수를 가지는 분류기의 화행을 선택하여 분

[†] 이 논문은 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. NRF-2013R1A1A2009937)

[‡] 학생회원 : 동아대학교 컴퓨터공학과
nh.song.89@gmail.com

[§] 학생회원 : 동아대학교 컴퓨터공학과
kmbae0722@gmail.com

[¶] 종신회원 : 동아대학교 컴퓨터공학과 교수
youngjoong.ko@gmail.com

논문접수 : 2015년 05월 08일

심사완료 : 2015년 10월 14일

류를 진행한다.

하지만 이러한 방법은 각 화행 분류기의 학습 데이터로 사용되는 화행별 발화의 수가 차이가 많이 나는 경우 발화의 수가 많은 화행으로 분류 결과가 편향(bias)되는 문제가 발생 할 수 있다[8]. 또한, 모든 화행 분류기에서 음의 분류 점수를 가지는 문제가 존재한다. 이 경우 상대적으로 양의 분류 점수를 가지는 경우보다 분류가 잘못될 확률이 높다.

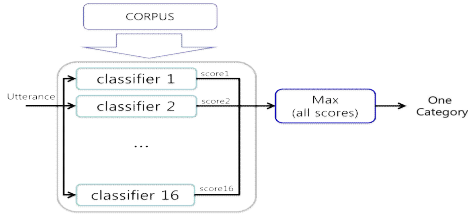


그림 1 일반적인 SVM 기반 다중분류기
(Fig. 1) General multi-classifier based on SVM

본 논문에서는 데이터 편향 문제를 해결하기 위해 발화의 수가 적은 화행 분류기에 분류 우선순위를 주고, 가장 먼저 양의 분류 점수를 가지는 분류기의 화행으로 분류를 진행함으로써 데이터 편향 문제를 해결하는 방법을 제안한다. 그리고 모든 분류기에서 음의 분류 점수를 가지는 경우 음의 분류 점수 상위 3개의 분류 모호성을 계산하여 모호성이 큰 경우 분류 결과가 틀릴 확률이 높다고 가정하고, 지지벡터기계의 분류 결과를 보정할 수 있는 보정 규칙을 이용하는 개선된 화행 분류 방법을 제안한다. 보정 규칙을 생성하기 위해서 본 논문에서는 변환기 기반 학습(Transformation-based Learning, TBL)을 이용한다. 본 논문의 구성은 다음과 같다. 2절에서는 관련 연구에 대해 살펴보고, 3절에서 본 논문에서 제안하는 방법들을 살펴본다. 4절에서는 제안한 방법을 실험을 통해 평가하고, 마지막으로 5절에서는 결론을 낸다.

2. 관련 연구

화행 분류를 위한 초기 연구들은 화행 태그가 부착된 학습 데이터로부터 규칙을 추출하는 방법을 이용하였다 [9]. [9]는 학습 데이터로부터 언어학적인 규칙과 대화 흐름도를 추출하여, 두 단계로 화행을 분류하는 시스템을 제안하였다. 최근에는 화행 태그가 부착된 학습 데이터를 기반으로 최대 엔트로피 모델(Maximum Entropy Model, MEM), 지지벡터기계(SVM)와 같은 분류 모델을 학습하는 기계학습(Machine Learning) 기반의 화행 분류 연구가 진행되고 있다. [3,4]는 구분 정보, n-gram 기반의 형태소-품사 정보, 품사 바이그램, 문맥(context) 정보를 학습 자질(feature)로 사용하여 화행 분류의 성능을 개선하는 방법을 제안하였고, [5]는 화행들을 두 개의 층을 가지는 계층 구조로 구분하고, 이 정보를 화행분류에 이용함으로써 화행 분류의 성능을 개선한 연구를 제안하였다. [6]은 현재 화행과 연관된 이전화행을 효과적으로 찾을 수 있는 담화 규칙을 기반으로 만들어진 담화 자질(discourse feature)을 추가하여 화행 분류의 성능을 개선한 연구를 진행하였으며, [7]은 각 화행의 확률 분포의 엔트로피를 사용하는 방법과 각 화행별 분포 정보를 활용

하여 학습 자질에 대한 가중치를 효과적으로 계산함으로써 화행 분류의 성능을 개선하는 방법을 제안하였다.

많은 선행 연구들은 지지벡터기계를 분류 모델로 사용하고 있으며, 분류 결과가 편향되는 문제와 모든 분류기에서 음의 분류 점수를 가지는 문제에 대해 고려하지 않고 있다. 본 논문에서는 지지벡터기계가 가지는 두 가지 문제를 해결함으로써 화행 분류의 성능을 향상시킬 수 있는 방법을 제안한다.

3. 제안하는 방법

3.1 분류 우선순위 적용 분류 모델

일반적으로 지지벡터기계 기반의 분류기에서 양수의 분류 점수를 가지면, 현재 분류기의 화행으로 분류될 확률이 높다. 하지만, 화행의 학습에 사용된 발화의 수가 상대적으로 적은 화행 분류기에서 추정된 양의 분류 점수는 많은 수의 발화로 학습한 분류기에 비해 낮은 양의 분류 점수를 가질 확률이 높기 때문에 정답 화행의 분류기에서 양의 분류 점수가 추정이 되어도, 만약 다른 화행 분류기에서 더 큰 양의 분류 점수를 가진다면, 분류가 잘못되는 문제가 발생한다. 예를 들어, 학습 발화의 수가 상대적으로 적은 Accept를 정답 화행으로 가지는 발화가 Accept 화행 분류기에서 양의 분류 점수를 가지더라도, 상대적으로 학습 발화의 수가 많은 Response 화행 분류기에서 가장 큰 양의 분류 점수를 가진다면, 이 발화는 Response 화행으로 잘못 분류가 된다.

본 논문에서는 학습 발화의 수가 적은 화행의 분류기에 대해 분류 우선순위를 줌으로써 학습 발화의 수에 따른 분류 결과의 편향 문제를 해결하는 방법을 제안한다. 그림 2는 분류우선순위 적용 모델이다.

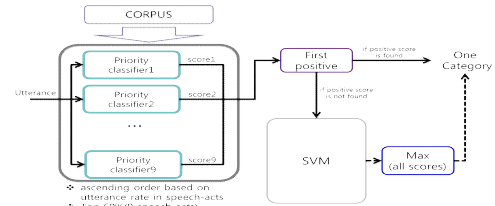


그림 2 우선순위 적용 모델의 개요
(Fig. 2) Overview of classification priority model

그림 2와 같이 학습 발화의 수가 적은 분류기부터 우선적으로 분류를 진행하고, 가장 처음 양의 분류 점수를 추정하는 분류기의 화행으로 분류한다. 실험을 통해 학습 발화의 수가 적은 순서부터 오름차순 정렬을 하여 상위 60%(9개의 화행)에 해당하는 분류기에 대해서 우선적으로 분류를 진행했을 때 가장 좋은 성능을 얻었다. 만약, 9개의 학습 발화의 수가 적은 분류기에서 양의 분류 점수를 가지는 분류기가 존재하지 않는 경우 분류 점수가 가장 큰 분류기의 화행으로 분류한다.

하지만, 이 방법은 분류기의 점수가 양인 경우에만 적용이 가능하다. 학습 발화의 수가 적은 화행의 발화는 모든 화행분류기에서 음의 분류 점수를 가지는 경우가 많다. 이 경우 일반적인 방법과 같이 가장 큰 점수의 발화로 분류를 할 경우 오답 화행으로 분류될 확률이 높다. 본 논문에서는 모든 분류기가 음의 분류 점수이면 발화에

대해 분류 결과가 오답일 확률이 높다고 가정하고, 이를 보정하는 규칙을 적용함으로써 화행 분류의 성능을 개선하는 모델을 제안한다.

3.2 후보정 규칙

모든 분류기에서 음의 분류 점수를 가지는 발화 중에는 분류 결과가 정답인 발화들도 존재한다. 본 논문에서는 분류 모호성 계산을 통해 분류결과가 틀릴 확률이 높은 발화를 구분하고, 구분된 발화에 대해서 보정 규칙을 적용한다. 그림 3은 지지벡터기계 기반의 분류 모델에 후보정 규칙을 적용한 최종 시스템이다.

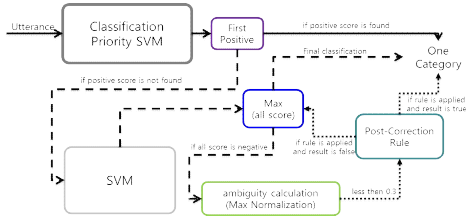


그림 3 제안하는 최종 시스템의 개요 (Fig. 3) Overview of proposed final model

3.2.1 분류 모호성 계산

모호성을 판단하기 위해 각 화행의 분류 점수가 모두 음수인 발화에 대해 각 분류기의 분류 점수를 순위화 한다. 음의 분류 점수를 가지는 화행의 경우 상위 3개의 분류 점수 결과가 비슷할 것이라는 가정 하에, 가장 큰 분류 점수와 두 번째, 세 번째 크기의 분류 점수의 차이를 수식 (1)을 통해 계산한다.

$$diff_i = (rank_1 - rank_2) + (rank_1 - rank_3) \quad (1)$$

- rank₁: 가장 큰 음의 분류 점수
- rank₂: 두 번째로 큰 음의 분류 점수
- rank₃: 세 번째로 큰 음의 분류 점수

$$F(diff_i) = \frac{diff_i}{Max(Train)} \quad (2)$$

$$Train = \{diff_1, diff_2, \dots, diff_n\}$$

수식 (1)을 통해 계산된 모호성 점수는 수식 (2)의 최대값을 이용한 정규화(max normalization)를 통해 0과 1 사이로 값을 변경한다. 상위 3개의 분류 점수가 비슷할수록 0에 가까운 값이 계산된다.

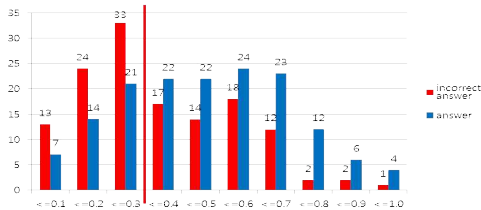


그림 4 분류 모호성 기반의 정답과 오답의 분포 (Fig. 4) The distribution of correct and incorrect answers based on classification ambiguity

그림 4는 분류 모호성 기반으로 정답 및 오답 분포이며, 분포를 통해 정답의 개수가 적으면서 오답의 개수가 많은 구간을 임계값(<math><0.3</math>)으로 한다. 임계값 이하인 발화에 대해서 분류 결과가 오답이라고 가정하고, 후보정

규칙을 적용한다.

3.3.2 후보정 후보 규칙 생성

후보정 규칙을 생성하기 위해서 후보 규칙을 생성하고, 가장 좋은 분류 결과를 얻을 수 있는 규칙들을 선정하는 작업을 진행한다. 이를 위해, 1990년대 Brill[10]에 의해 처음 소개되었으며, 품사 태깅, 문서 단위화, 개체명 인식 등 자연어 처리 분야에 널리 이용되고 있는 변환 기반 학습(TBL)을 이용한다.

본 논문에서는 3.2.1을 기반으로 하여 학습데이터로 학습하고, 학습데이터로 테스트(closed test)를 진행하여 잘못 분류되는 발화들을 추출하고, 추출된 발화에 대해 템플릿을 통해 후보 규칙을 생성한다. 템플릿에 사용되는 자질은 이전화행, 발화에 존재하는 단어, 형태소 태그를 이용한다. 발화에 존재하는 명사와 동사가 중요한 기준 자질이 될 것이라는 가정 하에 발화에 존재하는 명사와 동사의 태그를 기준으로 윈도우 사이즈가 5인 문맥 정보를 고려한 176개의 템플릿을 사용한다. 템플릿들은 연속적인 모든 단어 열을 고려한 형태와 단어 열을 고려하지 않은 형태를 사용하였다. 표 1은 추출된 발화에 적용할 템플릿들의 예를 보여준다.

표 1 템플릿 예 (Table 1) Template example

2-gram	$\{t_i, w_{i+1}\}, \{t_i, w_{i+2}\}, \{t_i, w_{i+3}\}, \{t_i, w_{i+4}\}$ $\{w_{i-1}, t_i\}, \{w_{i-2}, t_i\}, \{w_{i-3}, t_i\}, \{w_{i-4}, t_i\}$
3-gram	$\{t_i, w_{i+1}, w_{i+2}\}, \{w_{i-2}, w_{i-1}, t_i\}$
4-gram	$\{t_i, w_{i+1}, w_{i+2}, w_{i+3}\}, \{w_{i-3}, w_{i-2}, w_{i-1}, t_i\}$
5-gram	$\{t_i, w_{i+1}, w_{i+2}, w_{i+3}, w_{i+4}\}$ $\{w_{i-4}, w_{i-3}, w_{i-2}, w_{i-1}, t_i\}$

t 는 태그를 의미하고, w 는 단어를 의미한다. t 와 w 의 아래첨자는 index를 나타낸다. 2-gram 같은 경우 단어 열을 고려하지 않은 형태를 사용하고 나머지 3-gram, 4-gram 그리고 5-gram은 단어 열을 고려한 형태를 사용한다. 명사와 동사의 단어와 태그를 모두 사용하여 생성된 규칙은 명사와 동사의 태그만을 이용해 생성된 규칙보다 상세한 조건을 가지는 규칙이기 때문에 보정 규칙이 적용되지 않는 경우가 발생한다. 따라서, 템플릿의 기준 (t_i)으로 명사와 동사 태그만을 사용한다.

3.3.3 후보정 규칙 선정

후보 규칙을 생성한 후 가장 좋은 결과를 보여주는 규칙을 선정해야 한다. 이를 위해, 후보 규칙들의 적용 범위(coverage)와 성공률(success rate)을 고려한다. 적용 범위는 규칙이 적용되는 발화의 수를 의미하고, 성공률은 적용된 규칙 중 보정 결과가 정답인 경우의 비율을 나타낸다. 두 가지 조건을 계산하기 위해서 검증 테스트(validation test)를 이용하여 얻어진 데이터를 이용한다. 검증 테스트를 위해 학습 코퍼스를 학습 데이터와 검증 데이터로 구분하고, 학습 데이터로 분류기들을 학습한다. 학습된 분류기를 기반으로 검증 데이터에 대해 분류를 진행한 후 모든 분류기에서 분류 가중치가 음수인 발화들에 대해서 3.2.1을 이용하여 임계값이하인 발화들만을 추출한다. 추출된 발화들에 대해서 생성된 후보정 규칙들의 적용범위와 성공률을 계산 후 순위화를 통해 규칙을 선정한다. 실험을 통해 적용 범위는 1이상 성공률은 60%이상인

규칙들을 선정했을 때 가장 좋은 결과를 얻었다. 표 2는 후보정 규칙을 적용하여 오류를 수정한 예이다. 화행 정보 가 INFORM인 발화에 대해 지지벡터기계의 분류 결과는 오답인 ASK-REF이지만 후보정 규칙을 통해 정답인 INFORM으로 보정된 것을 볼 수 있다.

표 2 후보정 규칙을 적용한 오류 수정의 예
(Table 2) An example of error correction by the post-correction rule

Utterance		
EXPRESSIVE/PS 칠/NN 율/NBU 십/NN 사/NN 일날 /NBU 서울/NQE 로/JCA 돌아오/PV 르/EXM 예정이 /BE ㄴ테/EF		
Post-Correction Rule		
ASK-REF->INFORM : {NN ₀ , 사 ₁ , 일날 ₂ , 서울 ₃ , 로 ₄ }		
Answer Speech Act	SVM Result	Post-Correction Rule Result
INFORM	ASK-REF	INFORM

4. 실험 및 결과

4.1 실험 데이터

선행된 한국어 화행 분석 연구에서 사용된 대화 코퍼스(호텔, 비행기, 여행 예약)를 이용하였다[11,12,13,14].

표 3 각 화행별 학습 및 테스트 데이터 개수
(Table 3) Train and Test numbers in each speech act

speech act	Train	Test	speech act	Train	Test
Offer	33	8	Ask-if	450	101
Reject	88	22	Expressive	465	113
Suggest	167	37	Acknowledge	522	69
Accept	206	50	Opening	552	125
Promise	209	40	Introducing-oneself	553	141
Ask-confirm	243	82	Inform	974	250
Closing	278	69	Ask-Ref	1,120	257
Request	426	84	Response	2,060	484

대화 코퍼스는 10,285개의 발화와 17개의 화행으로 구성되었다. 이중 발화의 빈도가 너무 낮은 하나의 화행을 제거한 10,280개의 발화와 16개의 화행을 사용한다. 학습 데이터로는 8,348개의 발화를 그리고 테스트 데이터로는 1,932개의 발화를 사용하였고 표 3에 각 화행별 발화의 수를 나타낸다. Offer 화행과 Response 화행은 발화 수의 차이가 크고, 이는 학습된 분류기의 편향 문제가 발생할 수 있음을 보여준다.

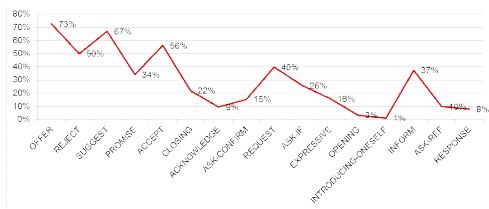


그림 5 화행별 음수 비율
(Fig. 5) The ratio of negative scores in each speech act

또한, 학습 발화의 수가 상대적으로 적은 발화들이 모든 분류기에서 음의 분류 점수를 가진다는 가정을 확인하

기 위해서 그림 5에서 보는 것과 같이 화행별로 분류기의 결과가 모두 음의 분류 점수로 나오는 비율을 검증 데이터를 통해 분석 하였다. 학습 발화의 수가 적은 화행들의 경우 상대적으로 모든 분류 점수가 음으로 나올 확률이 높다는 것을 알 수 있다.

4.2 자질 생성

선행 연구에서는 어휘자질과 문맥자질을 사용하고 있으며, 본 논문에서는 선행 연구와 비교를 위해 선행 연구에서 사용한 자질과 동일한 자질을 사용한다[5,6]. 표 4는 대화 코퍼스를 통해 추출된 자질의 예이다. 발화가 입력되면 형태소 분석 후 어휘자질로 유니그램 자질(어휘, 품사), 바이그램 자질(품사 바이그램)을 추출한다. 그리고, 문맥자질로 이전화행을 자질로 추출한다. '안녕하세요?'와 같이 대화를 시작하는 발화의 이전화행은 존재하지 않는다. 그렇기 때문에 이러한 발화의 이전화행을 'DIALOG-START'로 정의한다.

표 4 대화 코퍼스 예
(Table 4) Dialogue corpus example

Utterance	speech act
안녕하세요?	OPENING
대한항공입니다.	INTRODUCING-ONESELF
feature	
DIALOG-START/PS, 안녕하/PA, 시/EFP, 어요/EF, ?/S., PA_EFP, EFP_EF, EF_S.	
OPENING/PS, 대한항공/NQC, 이/JCP, 합니다/EF, ./S., NQC_JCP, JCP_EF, EF_S.	

4.3 실험 평가 방법 및 결과

본 논문에서는 학습 발화의 수가 적은 발화에 대해 분류 우선순위를 조정하여, 화행을 분류하는 방법과 분류 점수가 음수인 발화를 보정하는 규칙을 적용하는 방법을 제안하였다. 실험을 통해 제안한 방법에 대해 평가하였다. 표 5는 각 모델들의 성능 평가하기 위해 F1 결과를 나타내었다. 실험에서는 정확률(precision), 재현률(recall), F1 값을 사용하였다. Baseline은 선행 연구에서 사용한 어휘자질과 문맥자질을 학습 자질로 이용하여, 실험을 진행한 것이다.

표 5 각 모델들의 성능 평가 결과 (%)
(Table 5) Performance comparison of the three models (%)

speech act	F1-Measure		
	Baseline	Classification Priority Model	Final Model
micro	85.51	85.71 (+0.23, -1.59)	86.18 (+0.78, -5.38)
macro	75.51	75.76 (+0.33, -1.02)	76.45 (+1.24, -5.06)

학습 발화의 수가 적은 화행의 우선순위를 조정함으로써 Baseline보다 좋은 성능을 얻었으며, 후보정 규칙을 적용한 최종모델은 Baseline보다 상대적으로 0.78%의 성능을 향상시켰고, 5.38%의 에러 감소율을 보였다. 에러 감소율의 경우 상대적으로 향상된 0.78%에 Baseline의 에러율을 나눈 값이다.

표 6 선행 연구와의 성능 비교 (%)
(Table 6) Performance comparison with previous research (%)

speech act	F1-Measure				
	Our model	[5]	[6]		
Feature	(1)	(1)+(2)	(1)	(1)+(3)	(1)+(2)+(3)
micro	86.18	85.0	85.18	85.95	86.5

표 6은 선행연구와의 비교 결과이다. 표 6에서 (1)은 어휘자질과 문맥자질을 의미하고, (2)는 [5]가 제안한 두 개의 층을 가지는 계층 구조를 의미한다. (3)은 [6]가 제안한 담화자질을 나타낸다. 표 6에서 보는 것과 같이 제안한 방법은 같은 자질만을 이용했을 때 선행 연구인 [5,6] 보다 나은 성능을 얻었다. 다만, [6]에서 제안한 담화자질과 [5]에서 제안한 두 개의 층을 가지는 계층 구조를 모두 적용했을 때의 성능보다 제안한 방법이 낮은 성능을 보이고 있다. 향후, 제안하는 방법에 [5,6]에서 제안한 담화자질과 두 개의 층을 가지는 계층 구조를 적용한다면 더 나은 화행 분류 성능을 얻을 수 있을 것이라 기대된다.

표 7은 학습 발화의 수가 적은 순으로 상위 60%와 하위 40%의 화행들에 대한 성능 비교이다. 상위 60%의 화행들은 상대적으로 0.55%의 성능 향상과 1.61%의 에러 감소율을 보였고 보정 규칙을 추가함으로써 1.99%의 성능 향상과 5.83%의 에러 감소율을 보였다. 이는 하위 40%의 화행들 보다 좋은 성능을 보였으며, 이를 통해 제안된 화행 분류 방법이 학습 발화의 수가 적은 화행에 대한 분류 성능을 효과적으로 개선하는 것이라고 판단할 수 있다.

표 7 상위 60%와 하위 40%에 대한 화행 분류 성능 (%)
(Table 7) Performance of speech-act classification in top 60% and bottom 40% (%)

	F1-Measure		
	Baseline	Classification Priority Model	Final Model
Top 60%	65.86	66.22 (+0.55, -1.61)	67.17 (+1.99, -5.83)
Bottom 40%	87.92	88.02 (+0.11, -0.91)	88.39 (+0.53, -4.39)

5. 결론 및 향후 연구

본 연구는 지지벡터기계 기반의 화행 분류에서 발생하는 분류 결과의 편향 문제를 분류 우선순위를 조정함으로써 해결하고, 모든 분류 점수가 음수인 경우에 대한 분류 결과를 보정하는 규칙을 추가함으로써 분류 성능을 개선하는 방법에 대해 제안한다. 제안한 방법은 실험을 통해 선행 연구에서 사용한 자질을 동일하게 사용했을 때 선행 연구보다 더 나은 분류 성능을 보였으며, 향후 선행 연구에서 제안한 담화자질과 두 개의 층을 가지는 계층 구조를 적용한다면 더 나은 화행 분류 성능을 얻을 수 있을 것이라 기대된다.

참 고 문 헌

[1] Kyung-Sun Kim, Jung-Yun Seo, "Decision of the Korean Speech Act using Feature Selection Method", *Journal of Korean Institute of Information Scientists and Engineers*, Vol. 30, No. 3, pp.278-284, 2003.

[2] Min-Jeong Kim, Jae-Hyun Park, Sang-Bum Kim, Hae-Chang Rim, Do-Gil Lee, "A Comparative Study on Optimal Feature Identification and Combination for Korean Dialogue Act Classification", *Journal of Korean Institute of Information Scientists and Engineers*, Vol. 35, No. 11, pp. 681-691, 2008.

[3] Jongmin En, Songwook Lee, Jungyun Seo, "An analysis of Speech Acts for Korean Using Support Vector Machines", *The KIPS Transactions : Part B*, Vol. 12, No. 3, pp.365-368, 2005.

[4] Se-Jong Kim, Yong-Hun Lee, Jong-Hyeok Lee, "Korean Speech Act Tagging using Previous Sentence Features and Following Candidate Speech Acts", *Journal of Korean Institute of Information Scientists and Engineers*, Vol. 35, No. 6, pp.374-385, 2008.

[5] Sangwoo Kang, Youngjoong Ko, Jungyun Seo, "Hierarchical speech-act classification for discourse analysis", *Pattern Recognition Letters*, Vol. 34, pp.1119-1124.

[6] Kyungsun Kim, "Korean Speech Acts Classification System Using Improved Feature Extraction and Weighting Techniques", *doctorate thesis*, 2005.

[7] Youngjoong Ko, "New Feature Weighting Approaches for Speech-act Classification", *Pattern Recognition Letters, Elsevier Science*, Vol. 51, pp.107-111, 2015.

[8] A. Omuya, V. Prabhakaran, O. Rambow, "Improving the Quality of Minority Class Identification in Dialog Act Tagging", *Proc NAACL-HLT*, pp. 802-807, 2013.

[9] Won-Seug Choi, Jeong-Mi Cho, Jungyun Seo, "Analysis System of Speech Acts and Discourse Structures Using Maximum Entropy Model", *in the Proc. of 37th Annual Meeting of the ACL*, pp.230-237, 1999.

[10] E. Brill, "Transformation-Based Error-Driven Learning and Natural Language Processing: A Case Study in Part-of-Speech Tagging", *Proc. of the Computational Linguistics*, pp.543-565, 1995.

[11] Lee, Jae-won, Jungyun Seo, Gilchang Kim, "A dialogue analysis Model with statistical speech act processing for Dialogue Machine Translation", *Workshop in conjunction with (E)ACL' 97*, pp. 10-15, 1997.

[12] Hyunjung Lee, Jae-Won Lee, Jungyun Seo, "Speech Act Analysis Model of Korean Utterances for Automatic Dialog Translation", *Journal of Korean Institute of Information Scientists and Engineers*, Vol. 25, No. 10, pp. 1443-1452, 1998.

[13] Won Seug Choi, Jeong-Mi Cho, and Jungyun Seo, "Analysis System of Speech Acts and Discourse Structures Using Maximum Entropy Model", *In Proceedings of the 37th Annual Meeting of the Association for computational Linguistics*, pp. 230-237, 1999.

[14] Lee, Songwook, and Jungyun Seo, "Korean Speech Act Analysis Using Decision Tree", *In Proceedings*

*of the Conference on Hangul and Korean Language
Information Processing*, pp. 377-381, 1999.