

사용자 응답의 적절성 분류를 위한 대화 시스템 구성

정찬영¹, 이근배^{1,2}

¹포항공과대학교 인공지능대학원, ²포항공과대학교 컴퓨터공학과
{jcy1996, gblee}@postech.ac.kr

Configuring a dialogue system for classifying the appropriateness of user responses

Chanyoung Jung¹, Gary Geunbae Lee^{1,2}

¹Graduate School of Artificial Intelligence, Pohang University of Science and Technology

²Computer Science and Engineering, Pohang University of Science and Technology

요약

대화 시스템은 훈련 중 보지 못한 데이터를 입력받았을 경우 오작동의 가능성이 높으며, 특히 상용화 시에는 작은 오작동이라도 치명적인 결과를 낳을 수 있다. 따라서 시스템의 안정성을 확보하기 위해 우리는 대화 시스템에 사용자 입력의 적절성을 분류하는 기능을 탑재하였다. 해당 기능을 훈련시키기 위해서 우리는 슬롯-밸류쌍에서 슬롯을 교체하는 방법과 밸류를 교체하는 방법을 나누어 제시하였다. 슬롯을 무작위로 교체할 경우 발생하는 노이즈를 완화시키기 위해 새로운 추출 방법을 제시하였으며, 밸류를 교체할 때도 이 때 훈련시킨 모델을 이용하여 위음성의 가능성을 최소화하였다. 또한 전후방 철자 교정 모듈을 도입하여 사용자의 실수로 인한 모델의 오작동도 방지하고자 하였다. 우리는 다른 방법론과의 비교 실험과 절제 실험을 통해 각각의 방법론의 효용성을 입증하였다.

주제어: 대화 시스템, 자연어 분류, 사전 학습 모델

1. 서론

대화 시스템이란 인간과 대화할 수 있는 시스템을 일컫는다. 대화 시스템은 분야와 상관없이 자유롭게 대화할 수 있는 오픈도메인 시스템 [1, 2], 목적 수행을 위해 제한된 형태의 대화만이 가능한 목적 지향 시스템 [3, 4] 등 그 쓰임새와 기능이 다양하다. 특히 목적 지향 대화 시스템은 소비자 상담, 인터뷰 등 사용자로부터 얻고자 하는 정보의 양과 종류가 특정되어 있는 경우 매우 유용하다. 또한 그 용도가 제한적인만큼 비교적 오작동의 우려가 적어 상업적으로 사용되기에도 적합하다.

그러나 반대로 그 용도가 제한적인만큼 시스템의 학습 시 사용되는 데이터 또한 특정 도메인에 편향되어 있을 우려가 있다. 만약 훈련 시에 보지 못한 정보가 추론 과정에 입력되었을 경우 어떤 식으로 작동할지 예측 불가능하다. 따라서 목적 지향 대화 시스템에는 사용자의 응답의 적절성을 판단하고 분류할 수 있는 기능이 필요하다.

이전에도 대화 시스템의 입력을 분류하는 기능은 종종 사용되어 왔으나 대부분 다른 도메인의 데이터를 단순히 많이 모아놓고 이들을 음성, 실제 데이터를 양성으로 두어 별도의 분류기를 학습하는 방식에 불과하였다. 이 경우 오답 데이터의 수집 과정이 임의적인 만큼 위음성의 가능성이 존재한다. 이렇게 학습시킬 입력의 품질이 저하되면 출력의 품질 또한 보장할 수 없다는 것은 자명하다. 따라서 우리는 대화 시스템의 안정성을 확보하기 위해 더 정교하게 설계된 시스템 구조와 훈련 방법을 제안한다. 우리의 기여는 다음과 같다.

- 사용자 응답의 적절성을 분류하기 위해, 학습 데이터의 슬롯-밸류쌍에서 슬롯을 교체하는 방법과 밸류를 교체하는 방법을 각각 별개로 적용하였다. 슬롯-밸류 구조의 예시는 표 2에서 확인할 수 있다.
- 무작위 추출의 노이즈를 완화시키기 위해 새로운 추출 방법론을 제시하였다.
- 전후방 철자 교정기를 이용하여 사용자의 실수로 인한 시스템의 오작동 또한 방지하도록 하였다.

2. 방법론

여기서는 대화 시스템이 스스로 입력의 적절성을 분류하도록 훈련하는 방법론을 제시한다. 전체 시스템은 그림 1과 같이 대화 상태 추적기, 샘플러, 분류기, 철자 교정기의 4가지 모듈로 구성되어 있다. N 개의 모든 슬롯의 집합을 $S = s_1, s_2, \dots, s_N$ 라고 할 때, 대화의 t 번째 턴의 질문에 해당하는 슬롯을 s_t , 응답에 해당하는 밸류를 v_t , 해당 턴에 갱신할 신뢰 상태를 D_t 라고 한다.

2.1 대화 시스템

대화 시스템의 흐름은 규칙에 의존한다. 호텔 예약, AS 상담 등 필요한 정보가 한정적인 목적 지향 대화 시스템의 특성을 이용하여 정해진 순서에 의해 필요한 정보만을 사용자에게 질문한다.

신뢰 상태를 관리하는 대화 상태 추적기는 한국어로 학습시킨 대형 사전학습 언어모델, GPT2[5]이다. GPT2에 주어지는

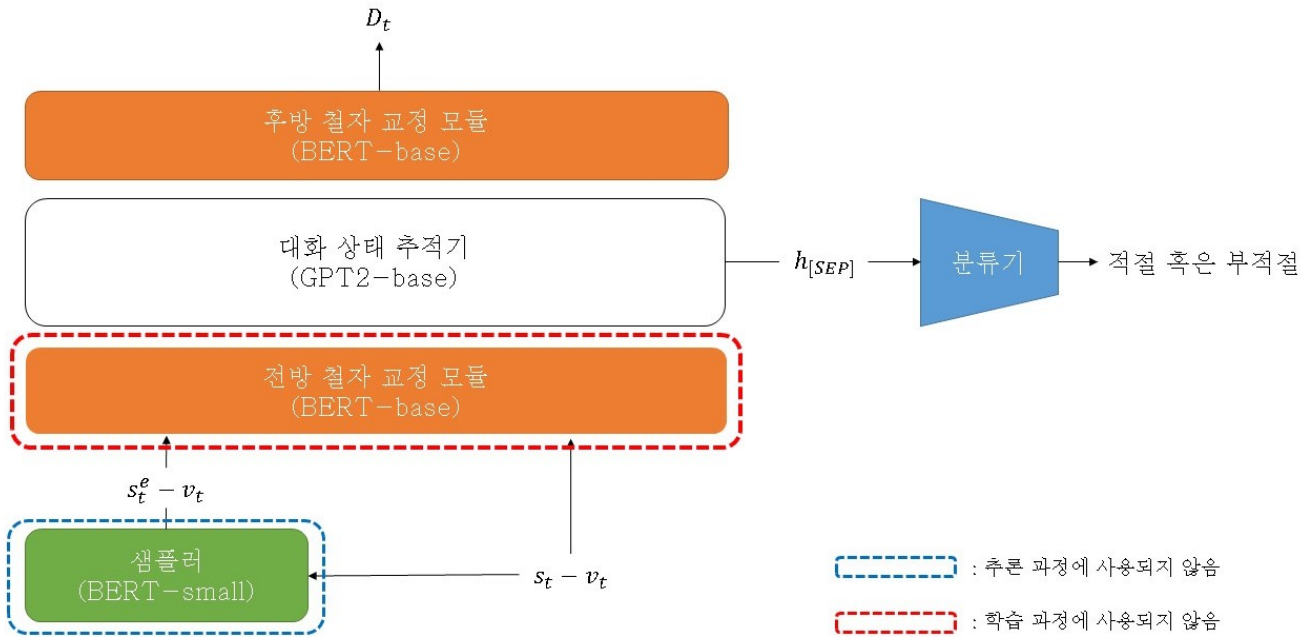


그림 1. 대화 시스템의 전체 구조. 실제 슬롯-밸류쌍이 입력으로 주어지면 BERT를 기반으로 한 샘플러가 슬롯을 교체하여 오답쌍으로 만든다. 실제 데이터와 조작된 오답 데이터가 분류기의 훈련을 위해 사용되며, 실제 데이터만이 해당 턴의 신뢰 상태 (Belief state) 를 갱신하기 위해 사용된다. 신뢰 상태는 대화가 진행됨에 따라 누적되는 정보를 의미한다. 대화 상태 추적기 (Dialogue state tracker) 의 전후방에 위치한 철자 교정 모듈은 모델 입력력의 품질을 개선시킨다. 적색 점선으로 둘러싸인 영역은 학습 과정에서 제외되며, 청색 점선으로 둘러싸인 영역은 추론 과정에서 제외된다.

입력의 형태는 다음과 같다.

[BOS] [ACTION] [REQUEST] S_t [/ACTION] V_t [SEP]
 [SLOT] S_t [/SLOT] [BELIEF] $S_t : V_t$ [/BELIEF] [EOS]

[SEP] 토큰은 GPT2의 사전학습에 사용된 특수한 토큰이다. 입력의 [SEP] 토큰의 순서에서 출력되는 은닉 상태를 $h_{[SEP]}$ 은 이후 분류기의 입력으로 사용된다. 이후의 학습 과정에서는 입력과 출력이 동일하도록 언어모델을 학습시키며, 추론 과정에서는 [BOS]부터 [/SLOT]까지의 토큰이 입력으로 주어졌을 때 [BELIEF] 이후의 내용을 생성한다. 생성된 내용을 바탕으로 D_t 를 갱신한다.

2.2 샘플러

대화 시스템을 훈련시키기 위한 데이터가 분류를 위한 잘못된 응답 데이터까지 포함하고 있는 경우는 흔치 않다. 따라서 대화 시스템이 사용자의 응답이 적절한지의 여부를 판단하도록 학습시키기 위해서는 음성 데이터를 보강하는 작업이 필수적이다. 우리는 대화 데이터의 슬롯-밸류쌍에서 슬롯을 교체하는 방법과 밸류를 교체하는 방법, 두 가지 방법으로 음성 데이터를 보강하였다.

2.2.1 슬롯 교체

슬롯 교체는 현재 턴의 슬롯 s_t 를, 이것을 제외한 집합에서 추출한 임의의 슬롯 s_t^e 로 교체하는 방법이다. 그 예시는 표 1에서 확인할 수 있다. 슬롯이 적절한 오답으로 교체되는 경우, 다른 슬롯에 매칭되는 밸류를 오답 데이터로 사용함으로써 주어진 데이터를 효율적으로 사용할 수 있다.

슬롯을 교체하는 알고리즘은 알고리즘 1에서 확인할 수 있다. 현재 턴의 슬롯 s_t 를 제외한 나머지 슬롯들을 $s_{t,1}, s_{t,2}, \dots, s_{t,N-1}$ 이라고 할 때, 이들을 모두 순회하며 현재 턴의 밸류 v_t 와의 매칭 점수를 계산한다.

	슬롯	밸류
원본 데이터	고소인 신분	대학생입니다.
적절한 슬롯 교체	날짜	대학생입니다.
부적절한 슬롯 교체	피고소인 신분	대학생입니다.

표 1. 슬롯 교체의 예시

매칭 점수를 계산하는 모델 F 은 한국어로 사전학습된 BERT를 사용하였다. '[CLS] 슬롯 [SEP] 밸류 [SEP]'의 입력

알고리즘 1 슬롯 교체 알고리즘**Input:** 현재 턴의 슬롯 s_t 와 밸류 v_t **Output:** 현재 턴의 슬롯을 교체할 오답 슬롯 s_t^e **Require:** 슬롯-밸류의 매칭 점수를 계산하는 모델 F 공집합 $C \leftarrow \emptyset$ 를 초기화**for** $n = 1, 2, \dots, N - 1$ **do** **if** $s_{t,n} = s_t$ **then** **break** 슬롯-밸류쌍의 매칭 점수 $c_{t,n} = F(s_{t,n}, v_t)$ 를 계산 $C \leftarrow C \cup (c_{t,n}, s_{t,n})$ c_t 가 낮은 K 개의 원소만을 추출 $C_K = \text{lowest}_K(C)$ $s_t^e = \text{random}(C_K)$

형태를 사용하여 실제 대화 데이터의 슬롯-밸류쌍을 이용해 정답 케이스를 만들고, 이것의 슬롯을 임의로 교체하여 오답 케이스를 설정하였다. 그 뒤 BERT를 이용한 분류를 수행할 때와 같이 [CLS] 토큰과 같은 위치의 은닉 상태를 선형 분류기에 입력시켜 슬롯이 교체되었는지 혹은 교체되지 않았는지를 판별하도록 하였다. 추론 시에는 0부터 1 사이 값으로 정규화된 분류기의 출력을 매칭 점수로 사용하였다.

여기서 슬롯의 교체는 무작위로 이루어져 다소 노이즈가 발생할 우려가 있다. 그러나 모든 슬롯과 현재 밸류 간의 매칭 점수를 완벽히 정렬하려는 것이 아니라 '적당한 오답'을 고르는 것이 목적이란 점에서 이 훈련 방법은 충분히 유의미하다. 잘 학습된 모델은 현재 밸류와 매칭되었을 때 위음성의 가능성이 없는 슬롯보다 그렇지 않은 슬롯에게 상대적으로 더 높은 매칭 점수를 주게 된다.

교체 후보인 $N - 1$ 개의 슬롯 중 모델이 출력한 매칭 점수가 가장 낮은 K 개의 후보만 남긴 뒤, K 개 중 하나를 무작위 추출한 결과가 최종적으로 교체할 오답 슬롯이 된다. 우리는 실험에서 $K=9$ 로 두었을 때 가장 좋은 결과를 얻을 수 있었다. 단순히 가장 낮은 점수의 슬롯으로 교체하는 것은 두 가지 측면에서 문제점이 있다. 하나는 앞서 언급했듯이 매칭 점수를 계산하는 모델이 노이즈로 인해 완벽한 정렬을 수행하지 못한다는 것이다. 두번째는 교체될 슬롯이 특정 슬롯에 편향되어 다양한 패턴을 학습하지 못할 수 있다는 것이다. 추출 방법이 성능과 다양성에 미치는 영향은 3.2에서 확인할 수 있다.

2.2.2 밸류 교체

목적 지향 대화 시스템에서의 사용자 응답은 도메인과 표현이 한정적이다. 따라서 전혀 보지 못한 응답이 입력되었을 경우 시스템이 어떻게 작동할지 예측 불가능하다. 따라서 우리는 도메인 외 문장을 다수 수집하여 슬롯-밸류쌍의 밸류를 교체하는 방법을 사용하였다.

방대한 양의 문장을 모은 뒤 그 중 임의로 추출하여 분류기를 학습시키는 시도는 예전부터 존재하였다. 그러나 단순히 무작위로 추출하여 음성 데이터를 조직하게 되면 위음성 데이터로 인해 성능이 저하될 가능성이 농후하다.

이를 해결하기 위해 2.2.1에서 언급했던 매칭 모델을 다시 한번 사용할 수 있다. 이 경우 현재 턴의 슬롯 s_t 를 그대로 대체로 현재 턴의 밸류 v_t 를 도메인 외 문장 v_t^e 로 교체한 뒤 매칭 모델에 입력한다. 이 때 매칭 점수가 0.5보다 낮으면 $s_t - v_t^e$ 쌍이 음성 케이스가 될 수 있으며, 반대의 경우 그렇지 않다고 볼 수 있다.

이렇게 도메인 외 데이터를 학습을 위해 추출할 때 매칭 점수를 이용하여 필터링하는 학습 방법론이 성능에 미치는 영향은 표 4에서 제시하였다.

2.3 철자 교정기

우리는 시스템의 입출력 품질을 개선하기 위하여 대화 상태 추적기의 전후방에 철자를 교정할 수 있는 모듈을 부착하였다. 두 모듈은 가중치를 공유하는 BERT[6]로 구성되어 있다. 마스크 토큰을 채우는 BERT의 훈련 방법을 문법 교정에 적용한 중국어 사례가 존재하나 [7, 8], 우리는 한국어에 두 단계에 걸친 교정을 적용했다는 점에서 차이가 있다. 또한 대화 상태 추적기와 단어장을 공유하지 않는다.

철자 교정기를 훈련시키기 위해서는 대화 데이터의 슬롯-밸류쌍에서 밸류만을 추출해야 한다. 이 문장들에 대해서 임의로 몇 음절을 가려놓으면 모델은 가려진 음절을 다시 복원하는 작업을 수행한다. BERT의 사전학습 방법과도 유사하나, 음절 단위의 노이즈를 추가한다는 점에서 차이가 있다.

철자가 교정되는 과정은 그림 2에서 볼 수 있다. 우선 대화 상태 추적기의 토큰라이저를 사용하여 사용자의 응답을 인코딩 후 바로 디코딩한다. 이후 단어장에 포함되지 않아 제대로 디코딩되지 못한 토큰을 미리 학습한 철자 교정기가 복원하도록 한다. 그리고 교정된 문장을 대화 상태 추적기에 입력한다.

대화 상태 추적기의 전방에만 철자 교정기가 존재할 경우 사용자의 과실에 의한 오타에 쉽게 대처할 수 있으나, 대화 상태 추적기의 단어장 자체가 부족한 경우 입력을 제대로 주더라도 출력 자체가 불가능하다. 따라서 대화 상태 추적기의



그림 2. 철자 교정 과정. 청색으로 표시된 음절은 사용자의 과실로 오타가 발생한 경우, 적색으로 표시된 음절은 단어장의 한계로 출력되지 못한 경우를 가정하였다. 우리의 시스템은 두 경우 모두 대처할 수 있도록 설계하였다.

후방에도 철자 교정기를 배치해 단어장 문제로부터 최소한의 안전망 역할을 수행하게 할 수 있다. 단, 대화 상태 추적기와 철자 교정기의 단어장이 모두 포함하지 않은 토큰이라면 당연히 출력되지 않겠지만 그럴 가능성은 희박하다.

결과적으로 이 방법론은 출력 문장의 품질 개선은 물론, 사용자 응답의 적절성 분류 성능에도 유의미하게 영향을 미친다.

그러나 사용자 응답 적절성을 판별하는 학습 과정에서 전방에 위치한 철자 교정기를 그대로 유지할 경우 성능이 저하될 우려가 있다. 2.2.2의 밸류 교체로 인해 전혀 다른 도메인의 문장이 오답 데이터로 주어졌을 때, 이것을 전방 철자 교정 모듈이 '정답처럼 보이게' 만들면 분류기는 정답 표현을 오답으로 학습하게 된다. 예를 들어 슬롯이 '피고소인 신분'일 때 '선생님 ㄹㅇㅇㅇ'와 같은 오답 문장으로 밸류가 교체되었다면, 철자 교정기는 이를 '선생님입니다'와 같이 고치게 되고 이후의 추론 과정에서 피고소인의 직업이 선생님인 경우 오답으로 판별하게 된다.

따라서 우리는 전체 시스템을 학습시킬 땐 전방에 위치한

슬롯	밸류
범죄 수법	선입금 사기를 당했습니다.
피고소인 신분	조건 만남을 하는 여자였습니다.
고소인 신분	저는 31살 회사원입니다.
날짜	2019년 10월 25일입니다.
...	...

표 2. 법률 대화 데이터 예시

철자 교정기를 제거하고, 추론 과정에서 다시 추가함으로써 이런 부작용을 억제하고자 하였다. 학습과 추론 과정을 분리하는 기법이 성능에 미치는 영향은 표 4에서 확인할 수 있다.

2.4 분류기

분류기는 단순한 선형층으로 구성하였다. 대화 상태 추적기가 입력을 [SEP] 토큰까지 인코딩했을 때 출력한 은닉 상태를 입력받아 현재 토큰의 슬롯-밸류쌍이 적절한가, 혹은 그렇지 않은가를 판별하는 이진 분류를 수행한다.

3. 실험

3.1 데이터

실험을 위한 데이터는 자체적으로 수집한 한국어 법률 대화 데이터를 사용하였다. 데이터의 예시는 표 2에 나와 있다. 대화 데이터는 총 4281개이며 도메인은 3개로 각각 성매매 유도 사기 (3084개), 중고거래 사기 (927개), 보이스피싱 (270개)이다. 성매매 유도 사기와 보이스피싱 도메인은 범죄 수법, 고소인 신분을 포함한 14개의 슬롯으로 구성되어 있다. 중고거래 사기 도메인은 거래 시스템, 실명 인증 여부 등 추가된 슬롯으로 인해 총 18개의 슬롯이 존재한다.

전체 데이터 중 3852개를 훈련용 데이터로, 429개를 검증 및 테스트용 데이터로 사용하였다. 또한 보이스피싱 데이터의 부족으로 인한 불균형 문제를 완화시키기 위해 기계번역기를 이용하여 한국어에서 영어, 영어에서 한국어로 번역하는 방법으로 보이스피싱 데이터를 853개 더 증강하였다. 또한 성능의 향상을 위해 전체 데이터에 대해 특정 음절을 제거하거나, 두 음절을 교체하는 증강 기법 [9] 을 사용하였다.

3.2 실험 1

우리는 2.2.1에서 무작위 추출로 인한 노이즈를 줄이기 위한 새로운 추출 방법을 제시하였다. 우리의 방법론의 효용성을 입증하기 위해 다른 추출 방법과 비교실험을 실시하였다. 실험에서 배치 사이즈는 8, 학습률은 3e-5로 두고 AdamW [10] 를 이용하여 최적화하였다. 최대 40에포크를 학습시키되 3

에포크 동안 성능이 향상되지 않으면 학습을 종료시켰다. 이 실험에서 단순히 분류 정확도 뿐만 아니라 모델이 얼마나 다양한 양상을 학습할 수 있는지도 측정하기 위해 엔트로피를 지표로 도입하였다. 엔트로피는 확률분포 Y 에 대하여 다음과 같이 계산한다.

$$Entropy(Y) = - \sum_{k=1}^K p(y_k) \log p(y_k) \quad (1)$$

따라서 다양한 표현이 추출되었을 경우 엔트로피가 높아지며, 그렇지 않을 경우 엔트로피는 낮아지게 된다.

각각의 비교 대상은 가장 매칭 점수가 높은 샘플을 고르는 방법 (max) 과 가장 낮은 샘플을 고르는 방법 (min) [11], 기준점에 근사한 샘플을 골라 학습을 적당히 어렵게 만드는 방법 (semi-hard) [12]이다.

세 가지 방법론 모두 무작위 추출의 한계를 극복하기 위해 제시되었다. 가장 매칭 점수가 높은 샘플을 고를 경우 정답 케이스와 유사한 어려운 오답 케이스가 생성된다. 따라서 학습이 어려운만큼 더 나은 일반화 성능을 기대할 수 있으나 위음성의 가능성이 높다. 반면 가장 매칭 점수가 낮은 샘플을 고를 경우 쉬운 오답이 생성되어 위음성의 가능성은 낮아지나 일반화가 어렵다. 따라서 매칭 점수가 기준점에 근사한 샘플을 고르는 방법은 학습을 위한 최적의 오답이 분포하는 위치를 직접 찾아내고자 한다.

	분류 정확도 ↑	엔트로피 ↑
무작위 추출	99.34	4.1669
max	97.30	3.4716
min	99.79	2.0706
semi-hard	98.66	3.5391
우리의 방법론	99.80	3.9561

표 3. 추출 방법론의 비교

실험의 결과는 표 3에 제시되어 있다. 공정한 평가를 위해 훈련 과정에서만 각기 다른 추출 방법을 사용하고, 추론 과정에서는 모두 무작위 추출을 사용하였다. 기준이 되는 무작위 추출의 경우 엔트로피가 가장 높아 다양한 슬롯이 오답으로 추출되었음을 알 수 있다. 그러나 분류 정확도로는 3번째로 높아 노이즈로 인한 성능 저하가 존재함을 알 수 있다.

가장 점수가 높은 슬롯을 추출하는 경우 위음성의 가능성이 극대화되어 가장 낮은 분류 정확도를 보인다. 반면 점수가 가장 낮은 슬롯을 추출하는 경우엔 안전한 오답만을 추출하기 때문에 정확도는 상대적으로 높아진다. 기준선을 지정하여 적

당히 어려운 오답을 생성하는 경우 딱 중간 정도의 정확도를 보인다. 세 경우 모두 추출을 특정 지점 근처에서 시행한다는 점에서 상대적으로 낮은 엔트로피를 보이며, 모델이 다양한 샘플을 학습하지 못하고 있다고 볼 수 있다.

우리의 방법론을 사용했을 때는 다른 방법론에 비해 엔트로피가 무작위 추출에 거의 근접하며, 분류 정확도도 가장 높게 측정되었다. 따라서 우리의 방법론이 모델에게 다양한 오답 표현을 습득할 수 있게 해준과 동시에 노이즈의 영향력을 성공적으로 완화하여 학습에 긍정적인 영향을 끼쳤다고 볼 수 있다.

3.3 실험 2

추출 방법 이외에도 제시한 방법론이 모델의 성능에 미치는 영향을 확인하기 위해 절제 실험을 실시하였다. 실험 세팅은 실험 1과 동일하게 유지하고 각 경우마다 하나의 방법론을 제거하였다. '학습-추론 분리'는 학습 과정에서 전방 철자 교정기를 제거하는 방법론, '도메인 외 데이터 필터'는 매칭 모델로 도메인 외 데이터를 안전하게 선별하는 방법론, '전후방 철자 교정 모듈'은 대화 상태 추적기의 전후방에 부착된 모듈을 의미한다. 실험의 결과는 표 4에서 확인할 수 있다.

	분류 정확도 ↑
전체 시스템	99.80
- 학습-추론 분리	99.77
- 도메인 외 데이터 필터	99.77
- 전후방 철자 교정 모듈	99.75

표 4. 절제 실험의 결과

각각의 방법론은 홀로 제거되었을 때 실험 1에서 추출 방법론을 교체했을 때만큼의 성능 저하가 발생하지 않았다. 따라서 대화 시스템이 사용자 응답의 적절성을 분류하는 과정에 가장 큰 영향을 미치는 것은 추출 방법론임을 알 수 있다. 그럼에도 불구하고 각각의 모듈은 처리하고자 하는 문제의 경우가 다르기 때문에 거의 모든 경우를 처리해야 하는 상용화 시스템의 경우 우리의 방법론의 도입이 충분히 유의미하다.

4. 결론

우리는 대화 시스템에서 사용자 응답의 적절성을 분류하기 위한 새로운 방법론을 소개하였다. 우리의 방법론은 비교적 활용하기 쉬운 간단한 방법들을 도입하여 해당 과제에 내재된 문제점을 해결하였다. 또한 적어도 우리의 경우에 있어서는 추출 방법론의 교체가 아주 중요한 역할을 하였다.

우리는 이 연구로 얻은 사실들을 통해 보편적으로 상용화될 수 있는 대화 시스템으로의 발전을 도모함과 동시에 본 연구가 후속 연구의 방향성을 제시하는데 도움이 되었으면 하는 바람이다.

감사의 글

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음 (IITP-2022-2020-0-01789). 또한, 이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2021-0-00575, 음성·텍스트 딥러닝 기반 보이스피싱 예방 기술 개발).

참고문헌

- [1] K. Shuster, E. M. Smith, D. Ju, and J. Weston, “Multi-modal open-domain dialogue,” *arXiv preprint arXiv:2010.01082*, 2020.
- [2] C. Xu, W. Wu, and Y. Wu, “Towards explainable and controllable open domain dialogue generation with dialogue acts,” *arXiv preprint arXiv:1807.07255*, 2018.
- [3] C.-S. Wu, A. Madotto, E. Hosseini-Asl, C. Xiong, R. Socher, and P. Fung, “Transferable multi-domain state generator for task-oriented dialogue systems,” *arXiv preprint arXiv:1905.08743*, 2019.
- [4] Z. Yan, N. Duan, P. Chen, M. Zhou, J. Zhou, and Z. Li, “Building task-oriented dialogue systems for on-line shopping,” *Thirty-first AAAI conference on artificial intelligence*, 2017.
- [5] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever *et al.*, “Language models are unsupervised multitask learners,” *OpenAI blog*, Vol. 1, No. 8, p. 9, 2019.
- [6] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” *arXiv preprint arXiv:1810.04805*, 2018.
- [7] Y. Hong, X. Yu, N. He, N. Liu, and J. Liu, “Faspell: A fast, adaptable, simple, powerful chinese spell checker based on dae-decoder paradigm,” *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT 2019)*, pp. 160–169, 2019.
- [8] S. Zhang, H. Huang, J. Liu, and H. Li, “Spelling error correction with soft-masked bert,” *arXiv preprint arXiv:2005.07421*, 2020.
- [9] J. Wei and K. Zou, “EDA: Easy data augmentation techniques for boosting performance on text classification tasks,” *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 6382–6388, Nov. 2019. [Online]. Available: <https://aclanthology.org/D19-1670>
- [10] I. Loshchilov and F. Hutter, “Decoupled weight decay regularization,” *arXiv preprint arXiv:1711.05101*, 2017.
- [11] J. Rao, H. He, and J. Lin, “Noise-contrastive estimation for answer selection with deep neural networks,” *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pp. 1913–1916, 2016.
- [12] J. Li, C. Tao, W. Wu, Y. Feng, D. Zhao, and R. Yan, “Sampling matters! an empirical study of negative sampling strategies for learning of matching models in retrieval-based dialogue systems,” *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 1291–1296, Nov. 2019. [Online]. Available: <https://aclanthology.org/D19-1128>