

Bi-LSTM 보조 신경망 모델을 결합한 생성형

한국어 Inverse Text Normalization 모델

조정제^o, 신동수, 조경빈, 한영섭, 전병기

LG유플러스

[jeong3080, ehdt89, jkb5570, yshan042, bkjeon]@lguplus.co.kr

Generative Korean Inverse Text Normalization Model

Combining a Bi-LSTM Auxiliary Model

Jeongje Jo^o, Dongsu Shin, Kyeongbin Jo, Youngsub Han, Byoungki Jeon

LG Uplus

요약

Inverse Text Normalization(ITN) 모델은 음성 인식(STT) 엔진의 중요한 후처리 영역 중 하나이며, STT 인식 결과의 가독성을 개선한다. 최근 ITN 모델에 심층신경망을 활용한 연구가 진행되고 있다. 심층 신경망을 사용하는 대부분의 선행 연구는 문장 내 변환이 필요한 부분에 토큰 태깅을 진행하는 방식이다. 그러나 이는 Out-of-vocabulary(OOV) 이슈가 있으며, 학습 데이터 구축 시 토큰 단위의 섬세한 태깅 작업이 필요하다는 한계점이 존재한다. 더불어 선행 연구에서는 STT 인식 결과를 그대로 사용하는데, 이는 띄어쓰기가 중요한 한국어 ITN 처리에 변환 성능을 보장할 수 없다. 본 연구에서는 BART 기반 생성 모델로 생성형 ITN 모델을 구축하였고, Bi-LSTM 기반 보조 신경망 모델을 결합하여 STT 인식 결과에 대한 고유명사 처리, 띄어쓰기 교정 기능을 보완한 모델을 제안한다. 또한 보조 신경망을 통해 생성 모델 처리 여부를 판단하여 평균 추론 속도를 개선하였다. 실험을 통해 두 모델의 각 정량 성능 지표에서 우수한 성능을 확인하였고 결과적으로 본 연구에서 제안하는 두 모델의 결합된 방법론의 효과성을 제시하였다.

주제어: 음성인식, 역정규화, 자연어 생성, STT, Inverse Text Normalization

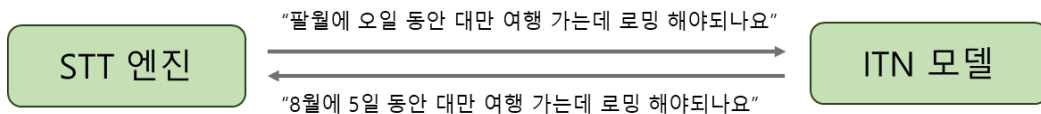


그림 1. STT 인식 결과에 대한 ITN 처리 예시

1. 서론

Inverse Text Normalization(ITN) 모델은 STT 인식 결과의 가독성 개선을 위해 철자, 숫자, 기호 등의 변환을 수행한다. 한국어 ITN 모델은 STT 엔진에서 인식된 결과에 대해 영문/숫자 변환을 하는 것이 대표적이다. ITN은 STT 엔진 인식 결과를 후처리하여 인식 결과의 가독성을 개선하고 사용자 경험을 향상시킨다. 또한 실 서비스에서 STT 엔진 인식

결과를 활용하는 개체명 인식, 의도 분류 등 자연어처리(NLP) 서비스 엔진의 성능 향상에 기여한다.

ITN 모델은 단일 언어 철자로 구성된 문장에서 변환이 필요한 부분을 인지하고 인지된 부분에 대한 변환을 진행한다. 문장에서 변환이 필요한 부분을 인지한다는 점에서 음성 합성 시스템(TTS)의 전처리 영역에서 활용되는 Text Normalization(TN) 모델보다 더 복잡한 연구 영역이다.

ITN 모델은 일반적으로 전통적인 규칙기반 방법[1, 3]과 최근 심층 신경망을 활용한 방법[4, 5, 6, 7, 8, 9]으로 연구가 진행되어왔다. 전통적 규칙기반 ITN 모델은 Weighted Finite-State Transducers(WFST)[2] 또는 정규표현식을 활용하여 사전에 텍스트 변환 규칙을 정의해 두고 이에 따라 변환 기능을 제공하는 방법론이다. 본 방법론은 모든 발화 케이스를 사전에 규칙화 할 수 없다는 점, 발화 문장의 전체 문맥을 파악할 수 없다는 점으로 한계점이 뚜렷하다. 심층 신경망을 활용한 ITN 모델은 대개 STT 엔진 인식 결과에 우선 신경망 기반 태깅 모델을 활용하여 변환이 필요한 부분을 토큰 단위 태깅을 하고, 다음 단계로 태깅된 결과에 대해 순차 대 순차(Sequence to Sequence) 기반 모델에 적용하거나 규칙기반 후처리 작업을 통해 ITN 처리 문장 생성한다. 본 방법론은 일반적으로 규칙기반 방법론과 비교하여 더 나은 성능을 보이지만 예러 디버깅이 힘들다는 점[4], 학습 데이터 구축 시 토큰 단위의 마이크로 한 데이터 태깅 작업이 필요하다는 점이 단점이다.

본 연구에서는 Bi-LSTM[10] 기반 보조 신경망을 활용한 생성형 한국어 ITN 모델을 제안한다. 생성형 ITN 모델은 실질적으로 ITN 기능 수행하는 BART[12] 구조의 언어 생성 모델이다. 생성 모델을 사용함으로써 BERT 등을 활용한 토큰 태깅 기반 접근 방법의 Out of vocabulary(OOV) 문제를 해결할 수 있다. 생성형 모델의 자기 회귀(Auto-regressive) 방식의 문장 생성은 빈번한 컴퓨팅 연산으로 서비스 추론 속도 측면에서 품질이 상대적으로 떨어질 수 있다. 이를 보완하기 위하여 Bi-LSTM 기반 보조 신경망을 활용하였다. Bi-LSTM 기반 보조 신경망은 STT 엔진 인식 문장의 ITN 처리 유무 판단, 고유명사 포함 유무 판단, 띄어쓰기 교정 기능을 병렬적으로 수행한다. BART 기반 생성 모델을 베이스로 구축한 한국어 ITN 모델과 Bi-LSTM 기반 보조 신경망 모델 출력 값에 따라 고유명사 처리, 띄어쓰기 교정 처리를 적용하고 처리 속도를 개선한 방법론을 제안한다.

2. 관련 연구

전통적으로 규칙 기반 ITN 모델은 Weighted Finite-State Transducer(WFST)[1,3] 활용한 방법론을 채택한다. 사전에 텍스트 변환 규칙을 정의해 두고 이에 따라 변환 기능을 제공하는 방법론이다. 이는 신경망 기반 접근법에 비해 컴퓨팅 연산 측면의 효율성을 갖는다. 하지만, 본 방법론은 규칙에 정의되지 않는 케이스에 대해 취약하다. 또한 문맥에 의거한 판단이 필요한 케이스에 대해 적절히 변환 작업을 진행할 수 없다.

이를 보완하여 최근 신경망을 활용한 ITN 모델 선행 연구는 BERT[14] 구조의 모델을 활용하여 STT 인식 결과 토큰에 대해 태그를 분류하고 분류된 태그를 기반으로 규칙기반 후처리 또는 다른 신경망 기반 후처리 방법을 통해 ITN 문장을 생성하는 방식으로 717 -

이루어진다. 하지만, 인식 토큰에 대한 태그 값을 분류하는 방법론은 한국어 대상 ITN에는 부적합하다. 영어권 문장 대상으로는 숫자/기호 변환이 주로 이루어 지지만, 한국어는 숫자/기호를 포함한 영문 변환까지 필요하다는 점에서 학습 시 토큰 별 분류 태그의 수에 비례하는 높은 컴퓨팅 자원이 요구된다. OOV에 취약하다는 단점이 존재한다. 최근에는 규칙 기반 방법론과 신경망 기반 방법론을 결합한 연구[7]도 있었다. 그러나 해당 연구 역시 토큰 분류 접근법으로 신경망을 활용했다는 점에서 한국어 ITN 모델에는 높은 컴퓨팅 연산이 요구되고 OOV 측면에서 한계점이 존재한다.

공통적으로 선행 연구에서는 STT 인식 결과를 그대로 사용한다는 점에서 띄어쓰기가 상대적으로 중요한 한국어의 특성을 고려하지 못하였다. 또한 ITN 처리가 필요 없는 인식 결과에 대해서도 불필요하게 ITN 모델 처리를 한다는 점, 명명된 그 자체가 정답인 고유명사 처리를 하지 못한다는 점은 실 모델 서비스 측면에서 공통적으로 한계점이 있다.

3. 제안 방법론

본 연구에서 제안하는 방법론은 Bi-LSTM 기반 보조 신경망 모델과 BART 구조 생성형 ITN 모델, 규칙기반 고유명사 치환 모듈을 결합하여 구성 되어있다.

3.1 Bi-LSTM 기반 보조 신경망

Bi-LSTM 기반 보조 신경망에서는 입력 문장 각 음절에 대한 임베딩 벡터를 활용하여 토큰, 음절 단위의 이진 분류를 수행하고 결과에 따라 3가지 기능을 제공한다. [BOS]의 분류 태그는 생성 모델 처리 필요 여부를 분류하고, [EOS]의 분류 태그는 고유 명사 포함 여부를 분류한다. 나머지 각 음절의 분류 태그 값은 음절 앞 분절의 포함 여부를 분류하여 STT 인식 결과의 띄어쓰기 교정에 활용한다.

ITN 처리가 필요 없는 문장은 생성형 모델 처리가 필요 하지 않다. [BOS] 토큰의 태그 값을 활용하여 생성형 모델 처리 필요 여부를 확인하고 불필요한 컴퓨팅 연산을 생략하여 전체 ITN 모델의 속도 성능을 향상시킬 수 있다.

또한, [EOS] 토큰의 태그 값으로 문장 내 고유명사 포함 여부를 판단한다. 문맥 상 명명된 그 자체가 정답인 고유명사가 포함되어 있다고 판단된 경우 사용자 사전을 참조 후 정규 표현식을 활용하여 고유명사 변환을 진행한다.

마지막으로, 각 음절 별 분절 포함 여부에 따라 STT 인식 결과의 띄어쓰기 교정을 진행한다. 한국어 특성 상 띄어쓰기 여부에 따라 의미가 상이해 질 수 있다. 이는 생성형 모델의 입력으로 들어갈 STT 엔진 별 상이한 인식 결과에 대한 띄어쓰기 의존성을 없이 생성형 ITN 모델의 품질을 유지할 수 있다.

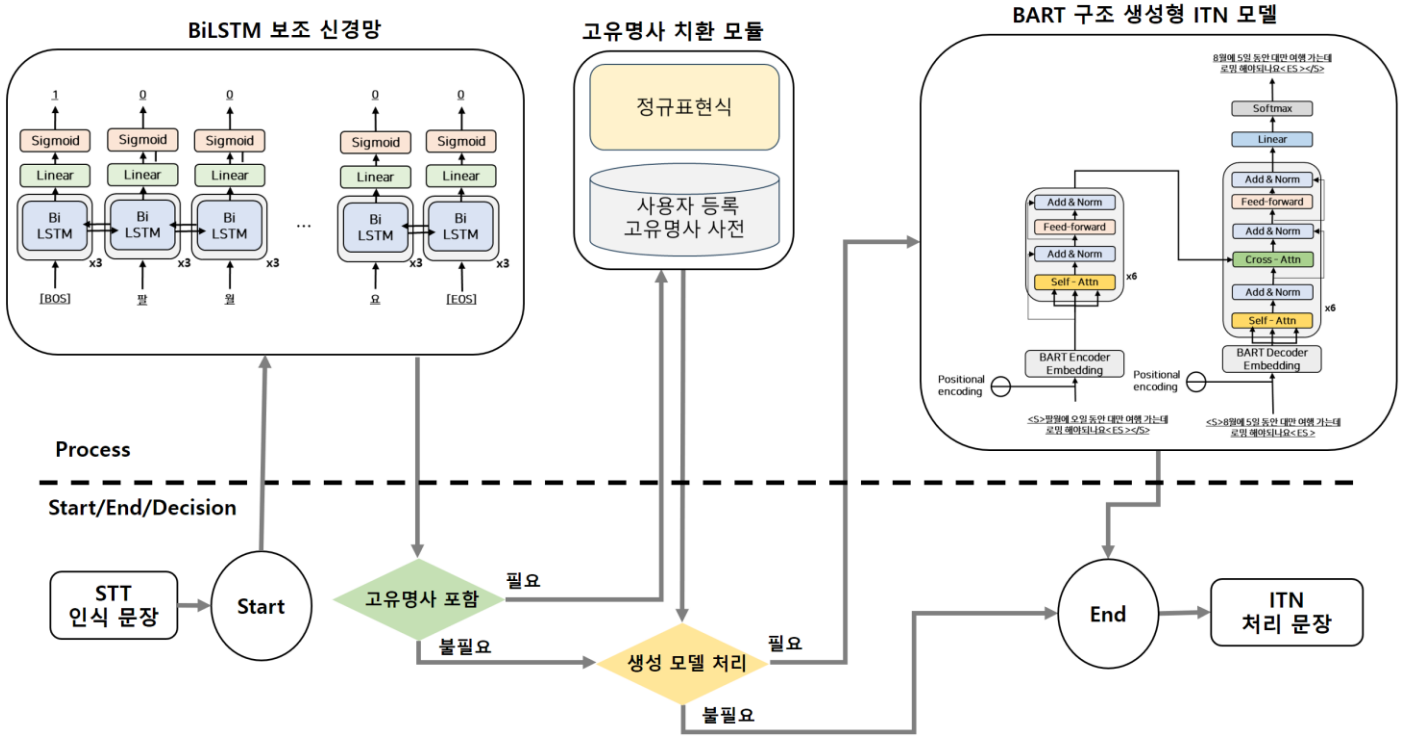


그림2. 제안하는 방법론의 구조도

3.2 BART 구조 생성 모델

본 연구에서 제안하는 생성형 모델은 트랜스포머의 인코더, 디코더를 사용하는 BART 구조를 갖는다. 생성형 모델은 입력 문장에 대해 순차 대 순차(Sequence to Sequence) 방식으로 ITN 처리된 문장을 출력한다. 기존 BART[12] Mike Lewis et al.가 제안한 모델 보다 어텐션 헤드의 개수, 은닉층 차원 개수를 줄여 BART-mini 사이즈로 정의하였다. 모델 구조 설계 후 별도 사전학습된 데이터의 영향을 우려하여 사전학습 없이 진행하였다. 가중치를 랜덤 초기화하고 ITN Task 학습을 진행하였다.

비해 학습 데이터 구축 시 원문과 변환 타겟 문장 병렬 코퍼스를 구축하면 된다는 점에서 용이성이 있다.

3.3 고유명사 치환 모듈

Bi-LSTM EOL 토큰의 태그 값에 따라 고유 명사 처리가 필요하다고 분류가 되면 고유명사 치환 모듈 처리가 진행된다. 문장 내에서 사용자 사전에 등록된 고유명사의 음절 순차에 일치하는 단어가 매치되면 정규표현식으로 단어 변환을 진행한다.

4. 실험

4.1 데이터셋 확보

제안 모델의 정량적 평가를 위해 실험 데이터 확보하였다. Bi-LSTM 보조 신경망과, BART 구조 생성형 ITN 모델은 각각 별도의 데이터셋으로 학습되고 평가되었다. Bi-LSTM 보조 신경망 학습에는 AI HUB에서 데이터 및 자사 데이터를 확보하였고 외부 맞춤형 검사 API를 활용하여 띄어쓰기 교정이 필요한 문장을 학습에서 제외한 6,305,008 건의 문장을 확보하였다. 고유명사 포함 여부에 대한 태그는 자사에서 보유하고 있는 개체명 인식 모델을 활용하였고, ITN 필요 여부에 대한 태그는 자사에서 보유하고 있는 Text Normalize(TN) 모델을 활용하였다. 띄어쓰기 교정을 위한 데이터는 각 음절 앞에 분절 유무에 따라 자체 태깅을 진행하였다.

	인코더/디코더 레이어 수	인코더/디코더 어텐션 헤드 수	인코더/디코더 은닉층 차원	파라미터 수
*BART mini	6 개	4 개	256dim	19M
BART base	6 개	12 개	768 dim	140M
BART large	12 개	16 개	1024 dim	400M

표1. BART 기반 생성형 ITN 모델 구성

BART mini는 모델 서비스 측면의 효율성을 높이기 위해 파라미터 개수를 대폭 줄였다. 인코더/디코더 레이어의 수 6개, 어텐션 헤드의 수 4개, 은닉층 차원 256dim으로 정의하였다. BART mini는 BART base에 비해 86.4%, BART large에 비해 95.2% 적은 파라미터의 개수를 갖는다.

기존 토큰 분류 방법론에 비해 OOV에 취약하지 않다. 또한 토큰 단위 태깅 작업이 필요한 기존 방법론에

예문	고유명사 포함 여부	ITN 필요 여부
<갤럭시 에스 십이 프로:content> 요금제 문의	TRUE	TRUE
<데이터 주교받기:content> 부가서비스 문의	TRUE	FALSE

낮 열두시 세명 예약	FALSE	TRUE
안녕하세요	FALSE	FALSE

표2. Bi-LSTM 모델 학습 데이터 태그 예시

BART 구조 생성형 ITN 모델의 학습에도 AI HUB 오픈 데이터 및 자사 데이터를 수집, 정제하여 10,975,829건의 병렬 코퍼스를 확보하였다.

원문	타겟
제 칠 회 지방 선거가 종료가 됐습니다	제 7 회 지방 선거가 종료가 됐습니다
십사 년 동안 쓰고 있는 거야	14 년 동안 쓰고 있는 거야
12 시 12 명 예약	12 시 12 명 예약
구십이 세잖아요.	92 세잖아요.

표3. BART 기반 생성형 ITN 모델 학습용 병렬 코퍼스 예시

4.2 손실함수

Bi-LSTM 보조 신경망의 손실함수는 Binary Cross-Entropy loss로 정의한다. 모델 내 선형 층(Linear layer)의 출력 값(S_i)과 시그모이드 함수를 거쳐 나온 각 토큰 별 확률 값 $f(S_i)$, Binary 인덱스의 산술식으로 계산된다.

$$f(S_i) = \frac{1}{1 + e^{-S_i}} \quad (1)$$

Binary Cross-Entropy loss

$$= - t_1 \log(f(S_1)) - (1 - t_1) \log(1 - f(S_1)) \quad (2)$$

마찬가지로 생성형 ITN 모델의 손실 함수는 Categorical Cross-Entropy loss로 정의한다. Categorical Cross-Entropy loss는 디코더의 선형 층(Linear layer)의 출력 값(S_i)과 소프트맥스 함수를 거쳐 나온 각 토큰 별 확률 값 $f(S_i)$, 후보 목표(target) 토큰(t_i) 인덱스의 산술식으로 계산된다.

$$f(S_i) = \frac{e^{S_i}}{\sum_j e^{S_j}} \quad (3)$$

Categorical Cross-Entropy loss

$$= - \sum_i t_i \log(f(S_i)) \quad (4)$$

위와 같이 각 모델의 손실 함수를 정의하고 손실 함수 출력 값을 최소화 하는 방향으로 신경망 파라미터를 최적화하였다.

4.3 학습

Bi-LSTM 보조 신경망과, BART 기반 생성형 ITN 모델은 각각 별도의 학습과 평가 과정을 거쳤다. 각 데이터 셋에서 Bi-LSTM 모델 10,000건, BART 모델 20,000 건의 데이터를 랜덤 추출하여 테스트 데이터셋으로 구성하였다. 검증 데이터셋 없이 나머지를 학습 데이터셋으로 구축하였다. Bi-LSTM 보조 신경망 학습 파라미터는 학습률 $2e^{-5}$, 옵티마이저는 Adam을 사용하였다. 평가는 음절 토큰 별 예측 값으로 F1 Score를 통해 평가하였다. 생성형 ITN 모델 학습 파라미터는 학습률 $2e^{-5}$, 옵티마이저는 AdamW, Float point16을 활용하여 GPU 활용을 효율화하였다. Bi-LSTM 보조 신경망 모델은 평가 데이터 손실 함수 값 수렴에 따라 20Epoch에서 학습을 종료하였다. 동일한 기준에 따라 BART 기반 생성 모델은 71Epoch에 학습을 종료하였다. 평가는 생성모델의 평가 지표 ROUGE-1, ROUGE-2, ROUGE-L 스코어와 예측 문장과 정답 문장의 완전 일치 여부를 평가하는 Accuracy를 활용하였다.

4.4 결과

평가용 데이터 10,000 건			
	Model	F1-Score	Accuracy
*Bi-LSTM	생성 모델 처리 필요 분류	0.936	0.962
	띄어쓰기 교정 부분	0.982	0.984
	고유명사 포함 여부 분류	0.919	0.856

표4. Bi-LSTM 모델 평가 결과

Bi-LSTM 모델의 평가는 생성 모델 처리 필요 분류, 띄어쓰기 교정 부분, 문장 내 고유명사 포함 여부 분류 3가지 TASK 별 F1-Score, Accuracy를 도출했다. 각 TASK 별 F1-score는 전부 90% 이상, Accuracy는 85% 이상의 성능을 확인했다.

평가용 데이터 20,000 건					
Model	Rouge-1	Rouge-2	Rouge-L	Accuracy	속도
*BART-mini	0.994	0.976	0.994	0.959	51ms
BART-base	0.996	0.977	0.997	0.961	349ms

표5. BART 기반 생성형 ITN 모델 평가 결과

BART 모델은 BART-mini와 BART-base를 비교 실험하였고, 두 모델 모두 사전 학습 없이 가중치 랜덤 초기화된 모델에서 실험을 진행하였다. 생성 모델의 평가 지표 Rouge score와 문장 완전 일치 여부를 평가하는 Accuracy 측면에서 두 모델의 성능에는 큰 차이가 없었다. 그러나 속도 측면에서 평균(한글 48음절 기준) BART-mini가 base 모델 대비 85% 빠른 성능을 보였다. 실험 환경은 CPU AMD EPYC 7713 Processor 환경에서 Pytorch[15]를 ONNX 표준 포맷으로 변환하였고 8bit 양자화[16] 하여 onnxruntime[16]으로 추론하였다.

예문	모델	고유명사 포함 여부 분류	생성 모델 필요 여부 분류	띄어쓰기 처리 결과	ITN 최종결과	처리 속도
피지컬백 재생해줘	LSTM +BART	TRUE	FALSE	피지컬 백 재생해줘	피지컬 100 재생해줘	2ms
	BART (w/o LSTM)	N/A			피지컬백 재생해줘	20ms
아이 피 티 비 요금제	LSTM +BART	FALSE	TRUE	아이피티비 요금제	IPTV 요금제	15ms
	BART (w/o LSTM)	N/A			아이 P T B 요금제	13ms
알뜰폰 요금제	LSTM +BART	FALSE	FALSE	알뜰폰 요금제	알뜰폰 요금제	2ms
	BART (w/o LSTM)	N/A			알뜰폰 요금제	21ms

표6. Bi-LSTM 보조 신경망 모델을 결합한 BART 기반 생성형 ITN 모델 추론 결과

5. 결론

본 논문에서 기존 ITN선행 연구의 단점을 보완한 생성형 ITN 모델을 제안하고, 한국어 띄어쓰기 등 보완 처리를 위한 보조 신경망 모델을 제안하였다.

AI-HUB 오픈 데이터 및 자사 내부 데이터를 수집, 가공하여 Bi-LSTM 기반 보조 신경망 모델과 BART기반 생성형 ITN모델을 구축하였다. 두 모델은 각 성능 지표에서 대부분 90% 이상의 높은 성능을 확인하였다. 두 모델을 결합하여 실 서비스에서 BART 기반 생성형 ITN모델의 취약점을 보완한 결과를 확인(참조: 표6)하였다.

본 논문에서 검증한 BART 기반 한국어 ITN 모델의 성능과 Bi-LSTM 기반 보조 신경망을 결합한 구조를 통해 STT시스템에서 중요한 후처리 영역으로서 한국어 ITN 모델 연구의 기초가 되기를 기대한다.

본 논문의 후속 연구로 본 모델에서 정의한 BART-mini 보다 파라미터 개수가 적은 모델로도 생성 모델의 성능을 유지할 수 있는지 실험을 진행할 예정이며, 성능을 유지하면서 서비스 속도 측면에서 빠른 모델을 확보할 수 있도록 기대한다.

참고문헌

[1] Shallow Fusion of Weighted Finite-State Transducer and Language Model for Text Normalization, Evelina Bakhturina et al, 2022, NVIDIA, <https://arxiv.org/abs/2203.15917>

[2] Weighted Finite-State Transducers in Speech Recognition, Mehryar Mohri et al, 2002, AT & T, Computer Speech & Language

[3] Weighted Automata in Text and Speech Processing, Mehryar Mohri et al, 2005, AT & T

[4] Thutmose Tagger: Single-pass neural model for Inverse Text Normalization, Alexandra Antonova et al, 2022, NVIDIA, <https://arxiv.org/abs/2208.00064>

[5] A Unified Transformer-based Framework for Duplex Text Normalization, Tuan Manh Lai et al, 2021, NVIDIA, <https://arxiv.org/abs/2108.098897>

[6] Inverse Text Normalization as a Labeling Problem, Siri Team, 2017, Apple Inc, <https://machinelearning.apple.com/research/inverse-text-normal>

[7] NEURAL INVERSE TEXT NORMALIZATION, Monica Sunkara et al, 2021, Amazon AWS AI, <https://arxiv.org/abs/2102.06380>

[8] Neural Models of Text Normalization for Speech Applications, Hao Zhang et al, 2019, Google Inc., Computational Linguistics

[9] RNN Approaches to Text Normalization: A Challenge, Richard Sproat et al, 2017, Google Inc., <https://arxiv.org/abs/1611.00068>

[10] Bidirectional recurrent neural networks, M. Schuster et al, 1997, IEEE Transactions on Signal Processing

[11] Long Short-Term Memory, S. Hochreiter et al, 1997, Neural Computation

[12] BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension, Mike Lewis et al, 2019, Association for Computational Linguistics

[13] Attention is all you need, Vaswano et al, 2017, NIPS

[14] <https://pytorch.org/>

[15] <https://onnxruntime.ai/>

부 록

A. 학습된 생성형 ITN 모델 추론 예시

원문: 근데 이 룡테일법칙 은 나온 게 이천사 년에 처음 나옵니다.
 예측: 근데 이 룡테일법칙 은 나온 게 2004 년에 처음 나옵니다.
 정답: 근데 이 룡테일법칙 은 나온 게 2004 년에 처음 나옵니다.

제35회 한글 및 한국어 정보처리 학술대회 논문집 (2023년)

원문: 칠억 오천 썩 일곱 개 해서 오십이억 오천 이고
우선지구 네 개 해서 십이억 을 했어요.
예측: 7억 5000 썩 7 개 해서 52억 5000 이고 우선지구 4 개
해서 12억을 했어요.
정답: 7억 5000 썩 7 개 해서 52억 5000 이고 우선지구 4 개
해서 12억을 했어요.

원문: 남은 문제 두 문제입니다.

예측: 남은 문제 2 문제입니다.

정답: 남은 문제 2 문제입니다.

원문: 팔십이 년예요?

예측: 82 년예요?

정답: 82 년예요?

원문: 한 번 한 번 두 번 한 번 더 셋 둘 셋 넷

예측: 1 번 1 번 2 번 1 번 더 3 2 3 4

정답: 1 번 1 번 2 번 1 번 더 셋 둘 셋 넷

원문: 일 위를 발표하는 게 맞겠죠 먼저 네 메이저 리그

예측: 1 위를 발표하는 게 맞겠죠 먼저 네 메이저 리그

정답: 1 위를 발표하는 게 맞겠죠 먼저 네 메이저 리그

원문: 이천일 년 베를린 영화제 은곰상 수상작인 이름 감독의
영화

예측: 2001 년 베를린 영화제 은곰상 수상작인 이름 감독의
영화

정답: 2001 년 베를린 영화제 은곰상 수상작인 이름 감독의
영화

원문: 네 보기 드리겠습니다 일 번 아이큐 이 번 인공지능 삼
번 지능지수.

예측: 네 보기 드리겠습니다 1 번 아이큐 2 번 인공지능 3 번
지능지수.

정답: 네 보기 드리겠습니다 1 번 아이큐 2 번 인공지능 3 번
지능지수.

원문: 저희 티비에스 티비에서 내일 오전 아홉시부터 이름
법무장관 후보자의 인사청문회 특집으로 저희가 생방송으로
나갈 예정인데요.

예측: 저희 TBS 티비에서 내일 오전 9시부터 이름 법무장관
후보자의 인사청문회
특집으로 저희가 생방송으로 나갈 예정인데요.

정답: 저희 TBS 티비에서 내일 오전 9시부터 이름 법무장관
후보자의 인사청문회
특집으로 저희가 생방송으로 나갈 예정인데요.

원문: 힌트를 선택할 때마다 십 점씩 감점이 있습니다.

예측: 힌트를 선택할 때마다 10 점씩 감점이 있습니다.

정답: 힌트를 선택할 때마다 10 점씩 감점이 있습니다.

원문: 한 분 계시고 이름 씨

예측: 1 분 계시고 이름 씨

정답: 1 분 계시고 이름 씨

원문: 스마트폰 티비에스 앱과 오십원에 유료 문자 샅에
영구오일 번으로 얼마지 정답
보내주시면 되겠습니다.

예측: 스마트폰 TBS 앱과 50 원에 유료 문자 샅에 0951 번으로
얼마지 정답 보내주시면
되겠습니다.

정답: 스마트폰 TBS 앱과 50 원에 유료 문자 샅에 0951 번으로
얼마지 정답 보내주시면
되겠습니다.

원문: 어~ 우리나라는 오이씨디 국가 중에서 탐정법이 없는
유일한 나라라고 방금 설명해 주셨는데요.

예측: 어~ 우리나라는 OECD 국가 중에서 탐정법이 없는
유일한 나라라고 방금 설명해
주셨는데요.

정답: 어~ 우리나라는 OECD 국가 중에서 탐정법이 없는
유일한 나라라고 방금 설명해
주셨는데요.

원문: 그렇게 정말 정신없이 변했던 지난 삼십 년이지만
변하지 않고 그대로 있는 곳이
그러니 오히려 더 신기하죠.

예측: 그렇게 정말 정신없이 변했던 지난 30 년이지만 변하지
않고 그대로 있는 곳이
그러니 오히려 더 신기하죠.

정답: 그렇게 정말 정신없이 변했던 지난 30 년이지만 변하지
않고 그대로 있는 곳이
그러니 오히려 더 신기하죠.

원문: 독서 추억 여행을 다녀왔습니다 이름 엠디였습니다
고맙습니다.

예측: 독서 추억 여행을 다녀왔습니다 이름 엠디였습니다
고맙습니다.

정답: 독서 추억 여행을 다녀왔습니다 이름 MD 었습니다
고맙습니다

원문: 한 모금이요

예측: 한 모금이요

정답: 1 모금이요

원문: 저녁에 시간 됴 피씨방서 경험치 이배 개꿀.

예측: 저녁에 시간 됴 PC 방서 경험치 2 배 개꿀.

정답: 저녁에 시간 됴 PC 방서 경험치 2 배 개꿀.