

미래문맥정보를 리용한 재귀신경망언어모형구축에 대한 연구

김철일, 리현순

경애하는 최고령도자 김정은동지께서는 다음과 같이 말씀하시였다.

《첨단과학기술분야에서 세계적경쟁력을 가진 기술들을 개발하기 위한 투쟁을 힘있게 벌려야 합니다.》

재귀신경망언어모형(RNNLM: Recurrent Neural network Language Model)[1-3]은 언어의 먼거리문맥특징정보들을 반영할수 있는 우월한 모형이다. 단어문맥을 제한하지 않고 임의의 길이를 가지는 문맥을 리력으로 사용하는 재귀신경망을 리용하여 언어모형을 추정하면 그 성능을 훨씬 더 개선할수 있다. 특히 두방향재귀신경망(Bi-Directional Recurrent Neural Network)언어모형은 단어의 리력문맥정보뿐만아니라 미래문맥정보도 동시에 모형화하는것으로 하여 언어모형의 정확도를 더욱 개선한다. 그러나 두방향재귀신경망에 의하여 미래문맥정보들을 모형화하는 경우 학습량이 방대해지고 N -최량(N -best)살창재득점에 의한 음성인식복호화에 리용하기 어렵다.

론문에서는 학습량이 적으면서도 미래문맥정보를 반영할수 있고 음성인식복호화에 리용하기 쉬운 재귀신경망언어모형구축방법을 제안하였다.

1. 재귀신경망에 의한 미래문맥정보의 리용

재귀신경망은 입력층 x 와 숨은층 s , 출력층 y 를 가진다. 입력층과 숨은층사이에는 무게행렬 U 와 W , 숨은층과 출력층사이에는 무게행렬 V 가 있다. 시각 t 에서 망의 입력은 $x(t)$ 이고 출력은 $y(t)$ 이며 $s(t)$ 는 숨은층의 상태이다.

재귀신경망의 구조를 그림 1에 보여주었다.

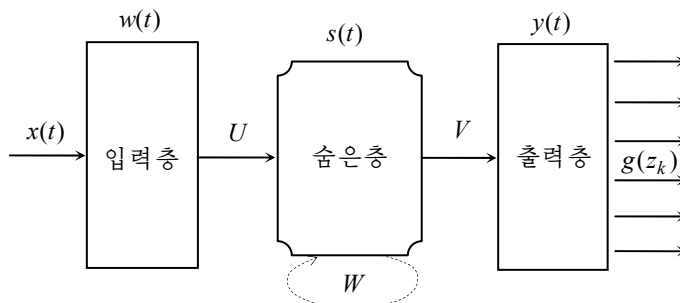


그림 1. 재귀신경망의 구조

숨은층과 출력층들은 다음과 같이 계산된다.

$$s_j(t) = f \left(\sum_i w_i(t) u_{ji} + \sum_l s_l(t-1) w_{jl} \right) \quad (1)$$

$$y_k(t) = g \left(\sum_j s_j(t) v_{kj} \right) \quad (2)$$

여기서 $f(z)$ 는 시그모이드활성함수, $g(z)$ 는 softmax 함수이다. 입력벡터 $x(t)$ 는 1-of- N 부호화를 리용하여 표현된 시각 t 에서의 단어 $w(t)$ 와 이전문맥층을 표현하는 상태벡터 $s(t-1)$ 로 이루어진다. 출력층 $y(t)$ 는 이전단어 $w(t)$ 와 문맥 $s(t-1)$ 이 주어진 조건에서 다음단어의 확률분포 $P(w(t+1)|w(t), s(t-1))$ 을 표현한다.

2. 재귀신경망에 의한 문맥단어정보의 리용

2개의 미래단어들을 리용한 재귀신경망언어모형학습체계의 구조를 그림 2에 보여주었다.

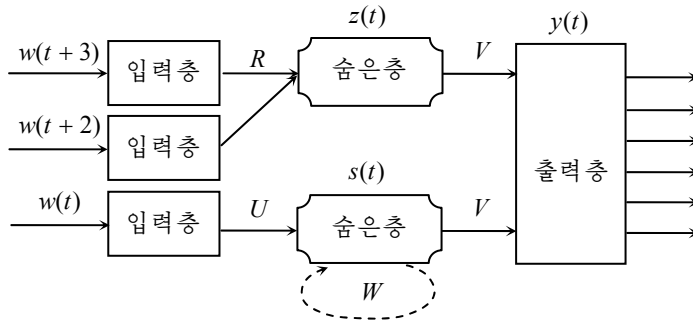


그림 2. 2개의 미래단어들을 리용한 재귀신경망언어모형학습체계의 구조

재귀신경망에서와 같은 방식으로 리력단어 w_1^t 들은 재귀신경망을 리용하여 모형화되며 유한개(k 개)의 미래단어 w_{t+2}^{t+k} 들은 앞방향결합신경망을 리용하여 모형화된다. 미래문맥단어개수가 0일 때 표준재귀신경망언어모형, ∞ 일 때 두방향재귀신경망언어모형과 같다.

미래문맥단어들의 개수를 2로 설정하였다. 출력층에서는 확률 $P(w_t | w_1^t, w_{t+2}^{t+k})$ 가 출력된다.

단어열 w_1^L 의 확률은 다음과 같다.

$$P_s(w_1^L) = \prod_{t=1}^L P(w_t | w_1^t, w_{t+2}^{t+k}) \quad (3)$$

3. 미래문맥정보를 리용한 재귀신경망언어모형학습방법

미래문맥정보를 리용한 재귀신경망의 한주기학습은 다음과 같이 진행된다.

- ① 시각 t 를 0으로 설정하고 숨은층에서 신경세포들의 상태 $s(t)$, $z(t)$ 를 초기화한다.
- ② t 를 1만큼 증가시킨다.
- ③ 입력층에 현재단어 w_t 와 미래단어 w_{t+2} , w_{t+3} 을 표현하는 단어벡터들을 입력한다.
- ④ 숨은층의 상태 $s(t-1)$ 을 입력층으로 복사한다.
- ⑤ $s(t)$, $z(t)$, $y(t)$ 를 계산한다.

- ⑥ 출력층에서 오차 $e(t)$ 의 그라디언트는 교차엔트로피를 리용하여 계산한다.
- ⑦ 오차를 역전파하고 무게들을 변화시킨다.
- ⑧ 모든 학습표본들이 처리되지 못했다면 단계 ②로 이행한다.

4. 성능 평가

실험에서는 논문에서 제안한 재귀신경망언어모형의 학습속도와 정확도평가를 진행하였다. 언어모형학습자료로 40M단어들로 구성되는 《로동신문》본문학습자료를 리용하였으며 이때 생성된 어휘크기는 60K이다. 평가자료로 언어모형학습에 리용되지 않은 10M의 소설본문학습자료를 리용하였다.

먼저 논문에서 제안한 모형의 미래단어개수에 따르는 학습속도평가를 진행하였다. 미래단어개수에 따르는 학습속도평가를 표 1에 보여주었다.

표 1. 미래단어개수에 따르는 학습속도평가

미래단어/수	0	1	2	5	∞
초당 학습단어/수	4.5K	4.5K	3.9K	3.8K	0.8K
분기/수	116.8	75.5	71.5	71.3	72.4

표 1에서 보여준 학습속도는 서로 다른 미래단어개수에 따르는 초당 학습단어개수와 분기수(PPLs)로 측정된 결과이다. 미래단어개수가 0일 때 이것은 표준재귀신경망언어모형 [1]과 같으며 미래단어들의 개수가 ∞ 로 설정되면 완전한 미래의 문맥정보를 가진 두방향 재귀신경망언어모형과 같다. 표 1로부터 미래단어개수를 증가시키에 따라 제안한 방법의 분기수는 감소된다는것을 알수 있다.

실험을 통하여 제안한 방법이 선행방법[1]과 거의 비슷한 학습속도를 가지며 두방향 재귀신경망언어모형보다 훨씬 빠르다는것을 확인하였다.

다음 제안한 방법을 조선어런속음성인식체계 《룡남산》에 적용하여 단어오유률평가를 진행하였다. 이때 대비를 위한 기준모형으로 저차원되돌이(backoff) 3-그람언어모형과 보간된 재귀신경망언어모형[2]을 리용하였다. 보간무게는 각각 0.75, 0.25로 설정하였다.

제안한 방법의 단어오유률을 표 2에 보여주었다.

표 2. 제안한 방법의 단어오유률

언어모형	미래단어/수	WER/%
선행방법[2]		9.2
	0	7.1
	1	6.8
제안한 방법	2	6.7
	5	6.6
	∞	6.4

미래단어들의 개수를 증가시키면 단어오유률이 감소된다는것을 확인하였다. 1개의 미래단어를 리용하면 일반재귀신경망에 비해 0.3%까지, 2개이상의 미래단어들을 리용하면 약 0.5%까지 단어오유률이 감소되었다.

실험을 통하여 미래단어문맥정보를 리용하여 학습된 재귀신경망언어모형은 일반재귀신경망언어모형에 비해 성능이 우월하다는것을 확인하였다.

맺 는 말

재귀신경망에 의하여 단어의 미래단어문맥정보를 모형화할수 있는 새로운 재귀신경망언어모형학습체계의 구조와 학습방법을 제안하였다. 미래단어문맥정보를 리용하는 재귀신경망언어모형은 일반재귀신경망언어모형에 비하여 계산량을 크게 늘이지 않으면서도 단어오유률과 분기수측면에서 우월하다.

참 고 문 헌

- [1] 김일성 종합대학학보(자연과학), 63, 4, 36, 주체106(2017).
- [2] Le, H-S et al.; ICASSP', 11, 5524, 2011.
- [3] Xunying Liu et al.; IEEE/ACM Transactions on Audio, Speech, and Language Processing, 24, 8, 1438, 2016.

주체109(2020)년 5월 5일 원고접수

Study of the Recurrent Neural Network Language Model Using Future Context Information

Kim Chong Il, Ri Hyon Sun

In this paper, we proposed the new recurrent neural network language modeling method that can model the future context information of the words, and evaluated its performance.

Keywords: language modeling, speech recognition, recurrent neural network