

조선어품사정보를 리용한 요소재귀신경망언어 모형구축에 대한 연구

리 현 순

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《과학기술정보사업을 강화하여야 합니다. 과학기술정보사업을 잘하여야 적은 밀천과 품을 들어 과학기술발전에 절실히 요구되는 귀중한 자료들을 얻을수 있습니다.》(《김정일선집》증보판 제15권 501페이지)

현시기 음성인식체계의 언어모형으로서 재귀신경망언어모형을 리용하기 위한 연구가 심화되어 이전의 N -그램언어모형[1]의 성능을 넘어서고있다. 재귀신경망학습[3]은 재귀적으로 련결된 숨은층을 리용하여 무한길이를 가지는 단어문맥을 모형화할수 있는 우월한 학습방법이다.

재귀신경망언어모형학습에서는 단어를 표현하는 1-of- N 지시벡토르로 된 하나의 요소정보(단어 그자체)만을 입력으로 리용하였다. 언어의 문법 및 의미정보와 같은 보충적인 언어정보들을 사용하지 않고 단어 그자체만을 학습에 리용하는 경우 문법적으로 류사한 단어들이 서로 련관이 없는 학습실투로 되기때문에 단어들사이의 련관관계를 정확히 반영할수 없게 된다.[2]

론문에서는 단어의 품사정보뿐만아니라 단어를 표현하는 여러가지 보충적인 정보들을 입력요소로 사용하여 재귀신경망언어모형의 정확도를 더욱 개선할수 있는 요소재귀신경망언어모형구축방법을 제안하였다.

1. 요소재귀신경망언어모형학습체계의 구조

요소재귀신경망언어모형학습체계의 구조를 그림 1에 보여주었다.

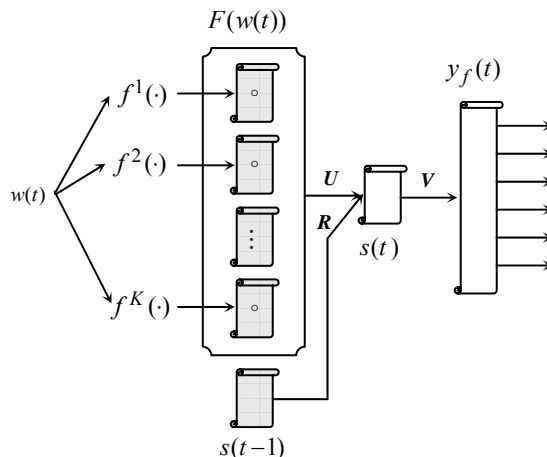


그림 1. 요소재귀신경망언어모형학습체계의 구조

요소재귀신경망의 출력층 $y_f(t)$ 는 단어 $w(t)$ 를 표현하는 K 개의 요소정보들과 리력 $s(t-1)$ 이 주어진 조건에서 단어 $w(t+1)$ 이 발생할 조건부확률분포 $P_f(w(t+1)|F(w(t)), s(t-1))$ 을 예측한다. 여기서 $P_f(\cdot)$ 은 요소재귀신경망언어모형확률분포를 의미한다.

$f_{w(t)}^k$ 는 단어 $w(t)$ 의 k 번째 요소정보를 표현하는 요소벡토르이며 $F(w(t))$ 는 단어 $w(t)$ 에 대한 K 개의 요소벡토르 $f_{w(t)}^k (k=1, \dots, K)$ 가 연결된 벡토르이다. 요소재귀신경망의 입력요소들로서 언어의 특성을 반영하는 임의의 정보들을 리용할 수 있다.

논문에서는 단어와 그것의 품사정보를 요소재귀신경망의 입력요소로 리용하였으며 품사정보는 조선어의 형태부단위를 기본으로 하여 음성인식에 합리적인 합성형태부단위들로 구성되어 있다. 음성언어의 특성에 맞게 설정된 96개의 조선어 품사정보를 리용하여 언어모형학습자료를 음성인식단위들로 분할하고 품사표식을 붙였다.

요소재귀신경망의 입력층에는 입력요소들이 1-of- N 부호화로 표현되어 입력된다. 단어 $w(t)$ 가 k 번째 요소정보 f^k 의 m 번째 성분을 나타낸다고 하면 요소벡토르는 벡토르의 m 번째 성분을 1로 하고 다른 모든 성분들을 0으로 하는 $|f^k|$ 차원벡토르 $f_{w(t)}^k$ 로 표시된다. 여기서 $|f^k|$ 은 k 번째 요소정보의 크기를 의미한다. K 개의 요소벡토르들을 연결하여 단어를 표현하는 요소벡토르 $F(w(t))$ 를 생성하고 요소재귀신경망언어모형의 입력으로 리용한다.

$f_{w(t)}^1$ 이 단어 그자체의 요소벡토르이고 $f_{w(t)}^k (k=2, \dots, K)$ 가 없는 경우 요소재귀신경망언어모형은 일반 재귀신경망언어모형으로 된다. 그러므로 재귀신경망언어모형은 요소재귀신경망언어모형의 한 경우로 볼 수 있다.

2. 요소재귀신경망언어모형의 학습

논문에서는 출력층분해된 재귀신경망언어모형학습[1]에 단어의 품사정보를 보충적인 입력요소로 리용하여 요소재귀신경망언어모형을 구축함으로써 재귀신경망언어모형의 정확도를 더욱 개선하였다.

출력층분해된 요소재귀신경망언어모형학습체계의 구조를 그림 2에 보여주었다.

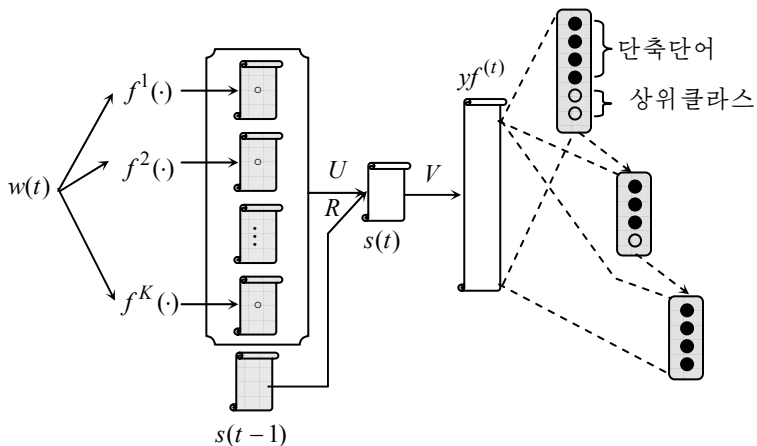


그림 2. 출력층분해된 요소재귀신경망언어모형학습체계의 구조

t 시각의 단어 $w(t)$ 에 대한 요소벡터 $F(w(t))$ 를 $x_i(t)$ 라고 하면 입력층에는 단어 $w(t)$ 의 요소벡터 $x_i(t)$ 와 이전 시각에서의 숨은층의 출력 $s(t-1)$ 이 입력된다.

숨은층에서의 계산은 다음과 같다.

$$s_j(t) = f \left(\sum_i x_i(t) u_{ji} + \sum_l s_l(t-1) r_{jl} \right), \forall j \in [1, H]$$

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

여기서 H 는 숨은층의 크기이며 u_{ji}, r_{jl} 은 무게행렬 U, R 의 성분들이다.

분해된 출력층에서의 단어 N -그램 확률은 다음과 같다.

$$P_f(w(t+1)|F(w(t)), s(t-1)) = P(c_1(w(t+1))|F(w(t)), s(t-1)) \times$$

$$\times \prod_{d=2}^D P(c_d(w(t+1))|F(w(t)), s(t-1), c_{1:d-1}(w(t+1)))$$
(2)

여기서 D 는 출력층분해를 위한 클래스나무의 깊이, $c_d(w_i)$ 는 단어 w_i 에 대응하는 클래스 또는 부분클래스, $c_D(w_i)$ 는 w_i 와 연관된 잎(단어)을 표현한다.

단축단어목록과 그밖의 단어들의 상위클래스들을 포함하는 출력층에서 확률분포는 다음과 같다.

$$P(c_1(w(t+1))|F(w(t)), s(t-1)) = g \left(\sum_j s_j(t) v_{mj} + \lambda_{im} \sum_{i=1}^C f_i(h, w_m) \right), \forall m \in [1, C]$$
(3)

여기서 C 는 첫째 출력층의 크기이고 $g(\cdot)$ 은 softmax 함수이다.

softmax 함수 $g(\cdot)$ 은 다음과 같이 표시된다.

$$g(z_d) = \frac{e^{z_d}}{\sum_x e^{z_x}}$$
(4)

예측되어야 하는 단어 $w(t+1)$ 이 포함되는 클래스나무의 경로에 속하는 나머지 부분 클래스층들에서의 확률분포는 다음과 같다.

$$P(c_d(w(t+1))|F(w(t)), s(t-1), c_{1:d-1}(w(t+1))) =$$

$$= g \left(\sum_j s_j(t) v_{mj} \right), \forall m \in [1, V'_d], \forall d \in [2, D]$$
(5)

여기서 V'_d 는 단어 $w(t+1)$ 이 포함되는 클래스나무의 d 번째 부분출력층의 크기이다.

3. 성능 평가

기준모형으로 출력층분해된 재귀신경망언어모형을 설정하고 학습속도와 수렴속도, 분기수, 단어오유률을 평가하였다.

요소재귀신경망언어모형의 성능평가를 표에 보여주었다.

표에서는 출력층분해된 재귀신경망을 dRNNLM, 요소재귀신경망언어모형을 fRNNLM

으로 표시하였다.

실험을 통하여 fRNNLM의 총체적인 학습시간이 dRNNLM에 비하여 3min정도 더 걸리지만 분기수와 단어오류률에서 개선이 있다는것을 확인하였다.

표. 요소재귀신경망언어모형의 성능평가

모형	한주기 학습시간/h	반복주기수	분기수	PPL	WER/%
dRNNLM	5	13	78.4	2.6	
fRNNLM	6.8	10	56.5	2.1	

맺는 말

조선어의 품사정보를 재귀신경망의 보충적인 입력요소로 리용하여 재귀신경망언어모형의 정확도를 더욱 개선할수 있는 요소재귀신경망언어모형 학습방법을 제안하고 그 효과성을 검증하였다. 그리고 재귀신경망언어모형의 학습시간을 크게 늘이지 않으면서도 성능이 높은 요소재귀신경망언어모형을 구축하였다.

참고 문헌

- [1] 김일성종합대학학보(자연과학), 63, 4, 36, 주체106(2017).
- [2] H. Schwenk; Computer, Speech & Language, 3, 21, 492, 2007.
- [3] L. H-S et al.; ICASSP'11, 5524, 2011.

주체107(2018)년 8월 5일 원고접수

Study on Factored Recurrent Neural Network Language Modeling by Using Information of Korean Parts of Speech

Ri Hyon Sun

In this paper, we enhanced the performance of neural network language model by using information of Korean parts of speech for training of recurrent neural network language model as additional input feature.

Key words: language modeling, recurrent neural network, speech recognition