



RNNs 종류

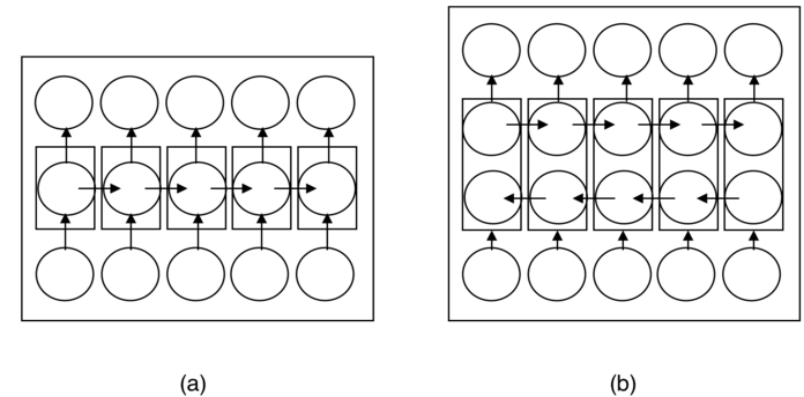
Tae-Won Lee,
leetee95@gmail.com
2020. 12. 15.

Bidirectional recurrent neural networks(BRNN)

(1997, 4387회 인용)

IEEE Transactions on Signal Processing, SCIE급, IF 5.230

- BRNN은 두가지 반대되는 방향의 전파가 있는 Hidden layer를 가짐.
- Output layer가 하나의 Input에 대해 두가지의 Input을 받을 수 있음.
- input값이 꼭 고정될 필요는 없음.



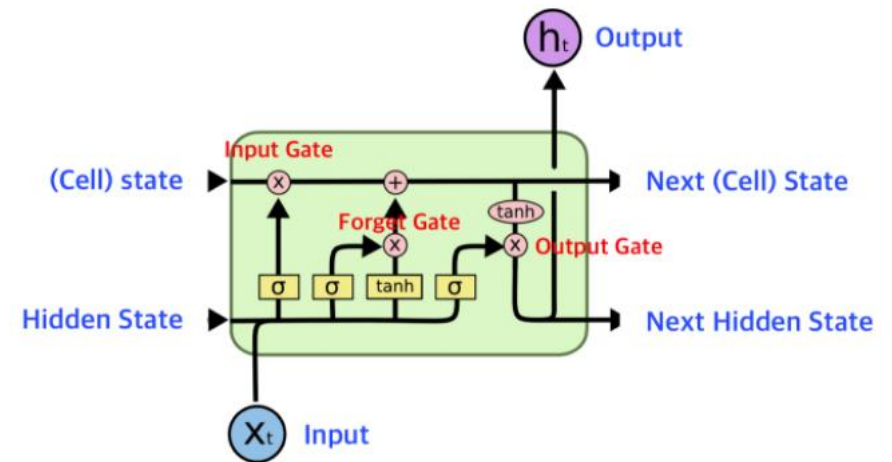
Structure overview
(a) unidirectional RNN
(b) bidirectional RNN

Long-short term memory(LSTM)

(1997, 40939회 인용)

Neural Computation, SCIE급, IF 2.207

- RNN의 문제점은 Gradient vanishing problem의 발생.
- 기존의 RNN에 Cell state가 추가됨.
 - Input gate : 과거 정보의 현재 Cell 에서의 중요도를 판단.
 - Forget gate : 현재 Input 값에서의 중요도를 판단.
 - Output gate : 지금까지 계산한 값 중 어떤 것을 내보낼지 결정.
- 기존의 RNN에 비해 Gradient vanishing 문제가 상당히 해소됨.

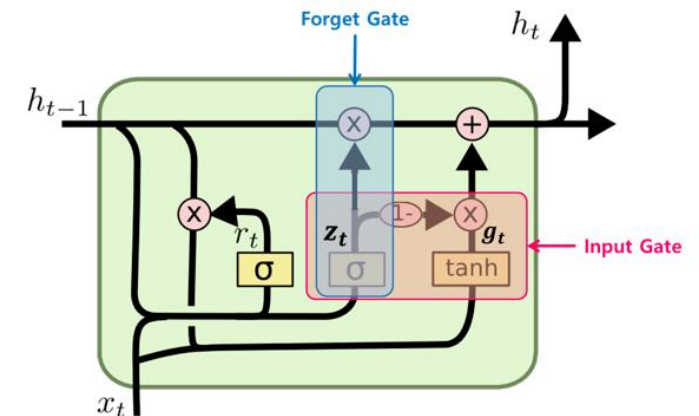


Gated Recurrent Units(GRU)

(2014, 5760회 인용)

Conference on Neural Information Processing Systems(NIPS), Qualis A1, 1987-현재

- Cell state가 존재하지 않음.
- 장기 의존성 문제에 대한 해결책을 유지하면서 Gate의 개수를 3개에서 2개로 줄어듬.
- 모든 연산에서 lstm보다 뛰어나다고 할 수 없음.



Minimal gated unit for recurrent neural networks

(2016, 151회 인용)

International Journal of Automation and Computing, ESCI급, IF 2.29

- 기존의 RNNs에서 많은 gate의 수와 복잡한 구조:
 - 시각화와 이해를 방해함.
 - gate수가 많다고 해서 좋은 결과가 보장되지 않음.
- 오직 하나의 gate만을 사용함.
- 학습 과정이 기존에 RNNs에 비해 빠름.

Hitnet: Hybrid ternary recurrent neural network

(2018, 21회 인용)

Conference on Neural Information Processing Systems, Qualis A1, 1987-현재

- Quantization은 neural networks에서 computational cost을 줄이기에 좋은 기술임.
 - CNNs에서 좋은 성과를 보임.
 - RNNs에서, 특히 low bit case에서 accuracy 감소를 겪음.
- 가중치 분포와 quantization 방식이 정확성의 차이를 만들어내는 것을 밝힘.
- HitNet은 $\{-1, 0, 1\}$ 과 같은 값들로 양자화를 진행했다.

unitary evolution recurrent neural networks

(2015, 424회 인용)

International Conference on Machine Learning, Qualis A1, 1980–현재

- 기존의 rnn에서 eigenvalues of Weight matrix가 절대값1에서 벗어날 경우에, vanishing and exploding gradients 문제가 발생.
- 대안으로, eigenvalues of absolute value 1를 가진 unitary weight matrix를 사용함.
 - 이경우는 complex(복소수) 영역의 hidden state를 고려할 때만 가능해진다.
 - 이 논문에서는 실제 구현을 real number를 이용했다.
- 제안된 모델의 성능은 기존의 lstm이나 gru를 능가했다.

Gated Orthogonal Recurrent Units: On Learning to Forget

(2019, 62회 인용)

Neural Computation, SCIE급, IF 2.207

- Unitary RNNs은 실제 사용에서 forget 능력이 부족했음.
 - GRU에서의 forget gate를 사용하여 이 문제를 해결.
- Complex number에서 gate를 사용하는 것에 문제가 생겨 orthogonal matrix에 한정됨.
- 제안된 모델의 성능은 장기 의존성에서 기존의 lstm, gru, Unitary RNNs 를 능가함.

Complex gated recurrent neural networks

(2018, 18회 인용)

Conference on Neural Information Processing Systems, Qualis A1, 1987-현재

- 시계열과 시퀀스 정보를 처리하는 데 널리 사용되는 RNN은 복소수 방식으로 큰 이익을 얻을 수 있음.
- Complex값, norm-preserving state를 GRU와 결합한 Hybrid 모델임.
- Norm이 보존되는 것은 rnn에서 시간이 지남에 따라 기울기를 보존하는것.
 - vanishing and exploding gradient problem에 효과적이다

Bayesian recurrent neural networks

(2017, 123회 인용)

Conference on Neural Information Processing Systems, Qualis A1, 1987-현재

- 이 논문에서 Bayesian방법을 훈련에 적용하여 RNN에 uncertainty and regularisation을 추가하는 방법을 검토함.
- Bayes by Backprop이 어떻게 적용되는지를 보여줌.
- 다른 maximum likelihood frameworks에서 Bayes by Backprop의 variance를 감소시키는 새로운 기술을 개발.

Independently Recurrent Neural Network(IndRNN)

(2018, 253회 인용)

Conference on Computer Vision and Pattern Recognition (CVPR), Qualis A1, 1987-현재

- Gate의 사용은 계산이 병렬이 아니라서, 계산의 복잡성을 증가시킴.
- 시간에 따른 상태가 병렬이 되기위해서 recurrent connection이 고정되어야함.
 - unitary evolution RNN에서 unitary recurrent weights를 정의하여 해결
- IndRNN은 뉴런들을 independent하게 만들고, 순환 가중치에 제약을 가함으로써 gradient 문제를 해결했다.

The Statistical Recurrent Unit(SRU)

(2017, 30회 인용)

International Conference on Machine Learning, Qualis A1, 1980–현재

- SRU는 Gate를 사용하지 않고, 통계적인 이동평균만을 이용하여 장기 의존성 문제를 학습함.
- Bayesian optimization scheme을 이용해 hyperparameters를 최적화 함으로써 lstm 및 GRU에 비교한 SRU의 성능을 보여줌.
- 이전의 Gated units보다 장기 의존성 문제에 대해 더 능숙함.

LSTM Fully Convolutional Networks for Time Series Classification

(2017, 300회 인용)

International Joint Conference on Artificial Intelligence, Qualis A1 , 1969~현재

- 과거의 머신러닝 방식을 이용한 다변량 시계열 자료의 분류는 더 많은 dataset의 전처리와 hyperparameters의 튜닝이 요구됨.
- .
- 이 논문에서는 데이터 전처리와 feature engineering을 요구하지 않는 두가지 딥러닝 모델을 이용하여 분류를 수행함.
- lstm과 fully convolutional network를 결합한 lstm-fcn, 그리고 attention LSTM을 이용함.

Recurrent neural network and a hybrid model for prediction of stock returns

(2015, 259회 인용)

Expert Systems with Applications, SCIE급, IF 5.89

- autoregressive moving average model, exponential smoothing model and rnn network를 혼합한 hybrid model임
- 최적의 가중치를 생성하는 방법으로 유전알고리즘을 사용함.
- 기존의 attention based encoder-decoder network and hierarchical attention networks들은 예측을 위한 것이 아닌 분류가 목적인 모형이었음.

Convolutional LSTM network: A machine learning approach for precipitation nowcasting (ConvLSTM)

(2015, 2670회 인용)

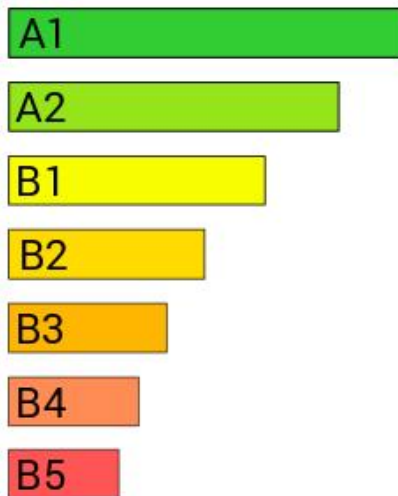
IEEE Conference on Computer Vision and Pattern Recognition, Qualis A1, 1985~현재

- ConvLSTM은 비교적 짧은 기간에 걸쳐 미래 강우 강도를 측정하는 precipitation nowcasting(강수량 예보) 모델.
- Input값과 prediction target값이 spatiotemporal(시공간) sequences
- 것은 혼란스러운 대기 상태 때문에 기존 머신러닝 방식으로 radar echo data로 모델을 만드는데 어려움이 있음. >> end-to-end 딥러닝 방식 사용

Appendix

Appendix

- Qualis



This conference ranking has been published by the Brazilian ministry of education and uses the H-index as performance measure for conferences. Based on the H-index percentiles, the conferences are grouped into performance classes that range from A1 (=best), A2, B1, ..., B5 (=worst).