

LSTM 실습

<https://youtu.be/rWq1ti9J24I>

Contents

주제

코드 설명



CryptoCraft LAB

주제

- LSTM을 이용한 주가 **예측 모델**
- LSTM : RNN의 vanishing gradient 문제를 해결한 신경망
- 10일치의 데이터를 이용해 다음 5일 후의 종가를 예측하는 모델



- 1일 후의 종가를 예측하지 않는 이유?

코드 설명

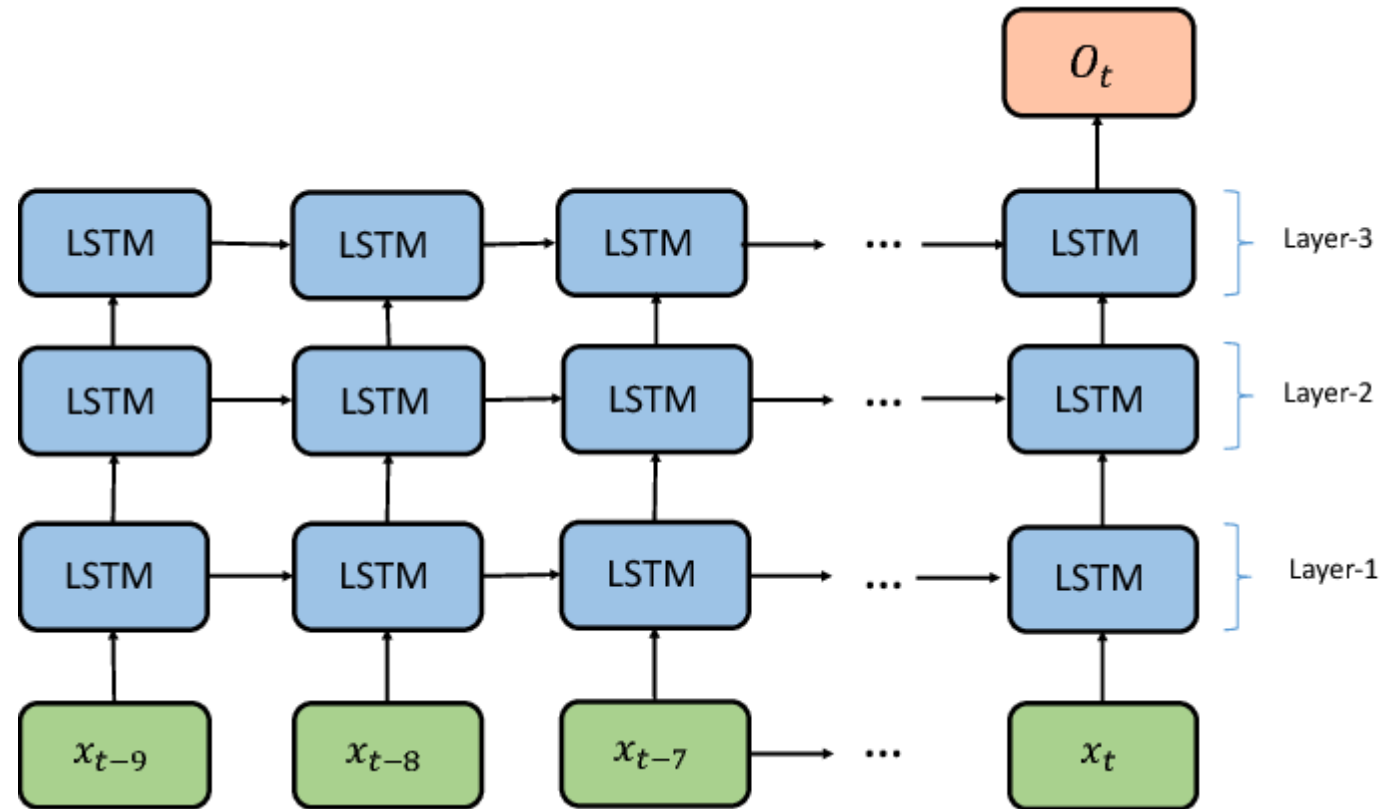
• Shallow / Deep copy

- `>>> a = [[1,2],[3,4]]`
- `>>> b = copy.copy(a)`
- `>>> a[1].append(5)`
- `>>> a`
- `[[1, 2], [3, 4, 5]]`
- `>>> b`
- `[[1, 2], [3, 4, 5]]`

- `>>> a = [[1,2],[3,4]]`
- `>>> b = copy.deepcopy(a)`
- `>>> a[1].append(5)`
- `>>> a`
- `[[1, 2], [3, 4, 5]]`
- `>>> b`
- `[[1, 2], [3, 4]]`

코드 설명

- **N_layers**
- Stacked LSTM



코드 설명

- Dropout

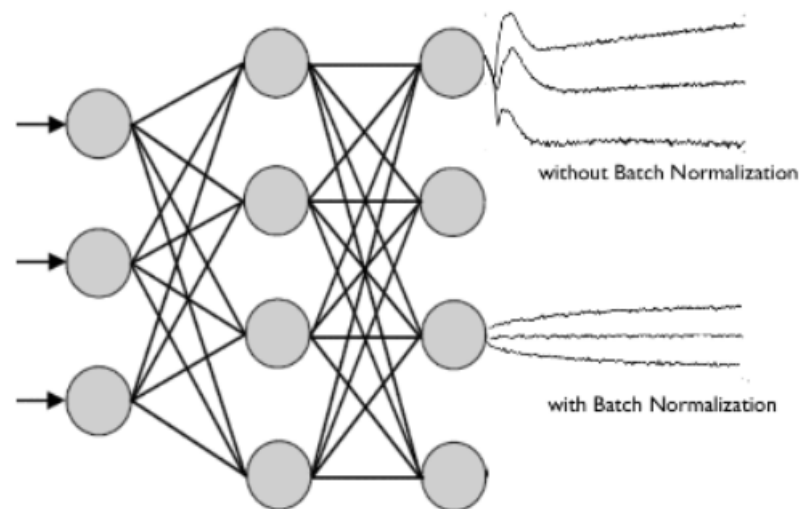
- 신경망의 뉴런을 부분적으로 생략하여 모델의 과적합을 해결해주는 방법

- train의 loss ↓ 그러나, test의 loss ↑



코드 설명

- **Use_bn**
- Bn : Batch normalization
- 신경망 안에 평균과 분산을 조정하는 과정이 포함되어 '**Internal Covariance Shift(내부 공변량 변화)**' 문제를 해결하는 방법
- 네트워크의 각 레이어나 Activation 마다 입력값의 분산이 달라지는 현상



코드 설명

- **Len(dataset)**
- Dataset = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
Len(dataset) = 10?

X_frames = 3

Y_frames = 1

=> $\text{Len}(\text{dataset}) - (\text{x_frames} + \text{y_frames}) + 1$

$10 - (3 + 1) + 1 = 7$

코드 설명

• `__getitem__`

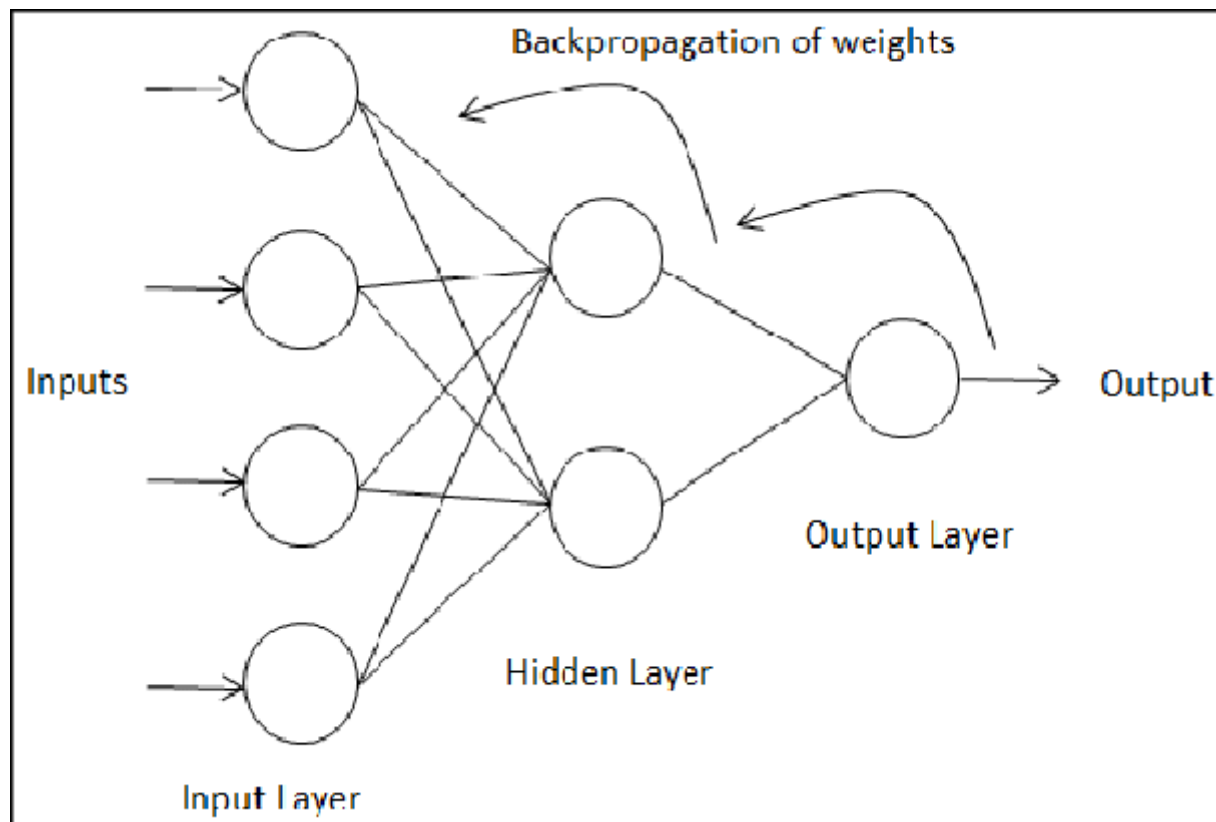
idx	X_frames										X_frames + y_frames				
	1일	2일	3일	4일	5일	6일	7일	8일	9일	10일	1일	2일	3일	4일	5일

- 1 $\log(1) - \log(2) \Rightarrow \log(1 / 2)$
- 1 $\log(1) - \log(2)$
- 2 $\log(2) - \log(2)$
- 2 $\log(2) - \log(2)$

- 10 $\log(10) - \log(20) \Rightarrow \log(10 / 20) = \log(1 / 2)$
- 10 $\log(10) - \log(20)$
- 20 $\log(20) - \log(20)$
- 20 $\log(20) - \log(20)$

코드 설명

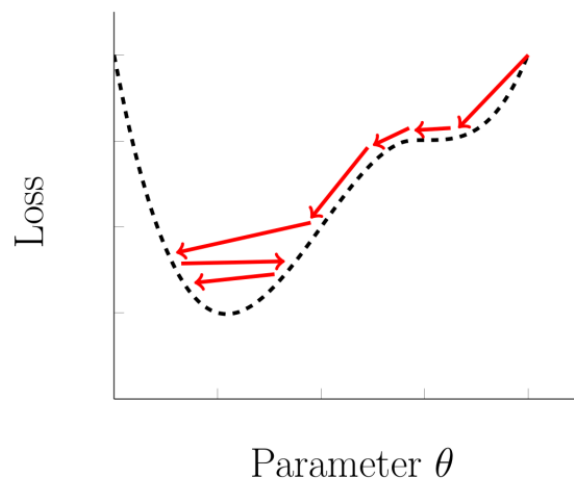
- **Back propagation**
- 역방향으로 진행하며
가중치를 갱신함
=> 효율적인 loss 계산



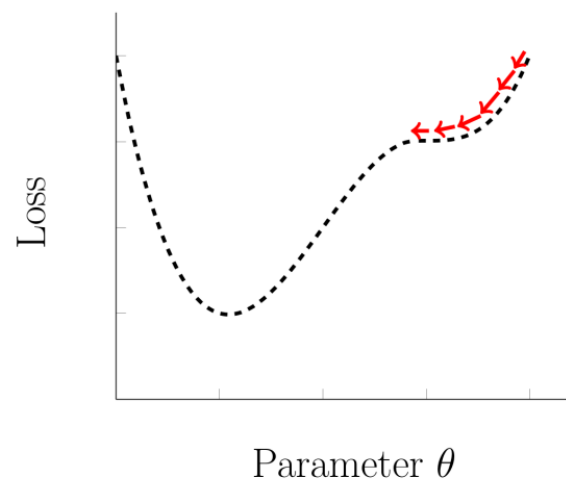
코드 설명

- **Optimization**
- loss가 가장 작은 (기울기가 0인) 곳을 찾기 위함

High Learning Rate



Low Learning Rate



Q & A

