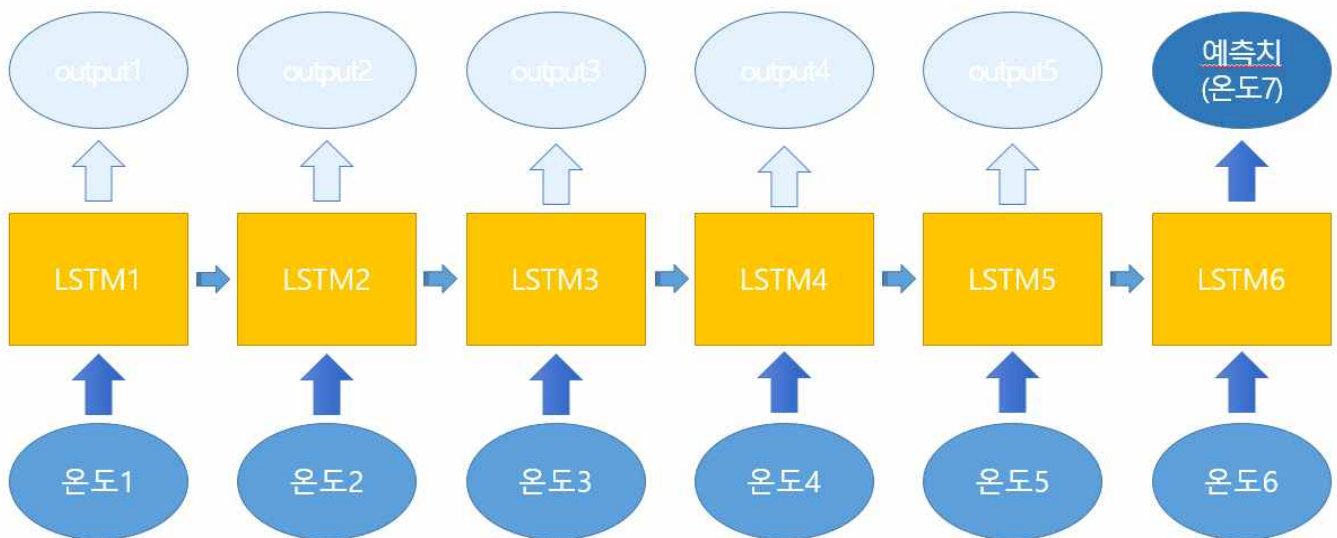


서울대학교 4차 산업혁명 아카데미  
빅데이터 플랫폼  
딥러닝(강유 교수님) 숙제3

RNN을 활용한 시계열(time-series)데이터 예측

이름 : 김 신

1. RNN DATA 구조 (Many-to-One)



2. 인자 탐색과정

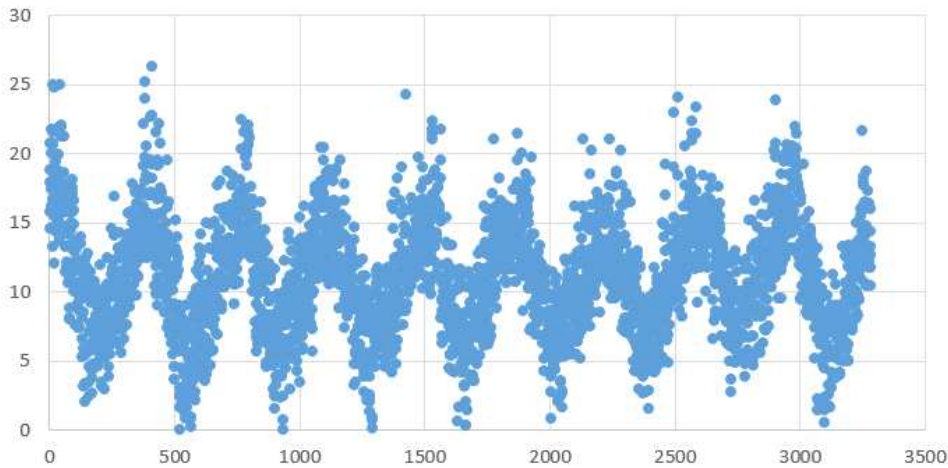
: 인자 탐색과정은 최초에 임의로 설정한 인자들을 기준으로 인자에 조금씩 변화를 줘가면서 RMSE 및 Valid Data Set의 그래프를 보면서 선택하였습니다.

※ 최초 설정 인자는 검정색 Bold, 최종 선택 인자는 파란색 Bold로 표시  
최초 설정 인자가 최종 선택 인자와 같을 경우는 파란색 Bold로 표시함

- Cell 쌓은 개수 = [1, 10, 20, 30, 40, 50]
- Cell 종류 = [Grub, LSTM]
- hidden\_dim = [32, 64, 128, 256, 512, 1024]
- learning\_rate = [0.1, 0.01, 0.001, 0.0001]
- iterations = [30, 60, 100, 200, 400, 800, 1600]
- Dropout = [0.6, 0.7, 0.8, 0.9, 1.0]

### 3. 데이터 조작 과정 및 결과

- Raw Data 탐색 : 계절별 특색이 있어, 여름과 겨울이 구분됨



- 데이터 조작 과정

- Normalization 및 Denormalization

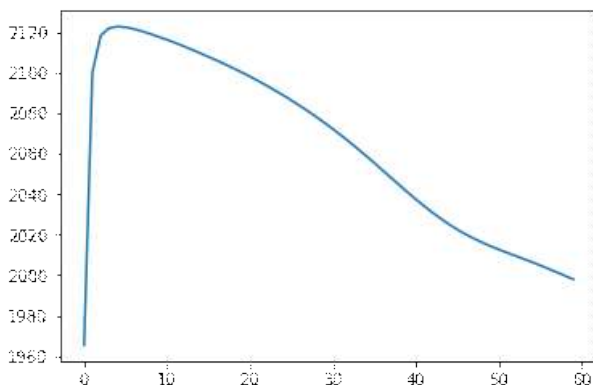
: Model을 학습 시킬 때는 Normalize한 Data를 집어넣은 후  
Model 평가 시에는 Denormalization한 Data로 평가하였음.

- 차원

- Input Data는 3차원 Output Data는 2차원.

: 처음에 차원을 잘 다루지 못하여 모델을 구동시키는데 어려움을 많이  
겪었습니다.

- Loss 그래프



Loss 그래프는 60까지 안정적으로 감소하는  
경향을 보여 Iteration을 60으로 결정하였음.

※ 60 이후엔 증가하는 등 불규칙적인  
특성을 보였음.

- 최종 결과

: 최적의 인자를 구하려고 다양한 인자들을 넣어 보았지만, RMSE를 줄여보려고  
하였지만, 좋은 결과를 얻지 못했습니다. RNN 구조에 대한 이해 부족 및 Data특성을  
잘 반영한 모델을 만들지 못하여 이러한 결과가 나온 것 같습니다.

- TRAIN\_RMSE : 3.3727521896362305
- TEST\_RMSE : 11.729297637939453

## ※ 건의사항

좋은 모델을 만들어 보려고 노력했지만, 특별한 인자를 탐색하지 않은 모델들보다도 나쁜 결과가 나왔습니다. 조교님께서 가능하시다면, 어떠한 점을 잘못하면 잘못된 모델을 만들게 되는지 이번에 다루어본 시계열 Data는 어떠한 방식으로 모델을 만들어야하는지 모든 학생들을 위한 전체적인 피드백을 해주신다면 학습하는데 큰 도움이 될 것 같습니다.