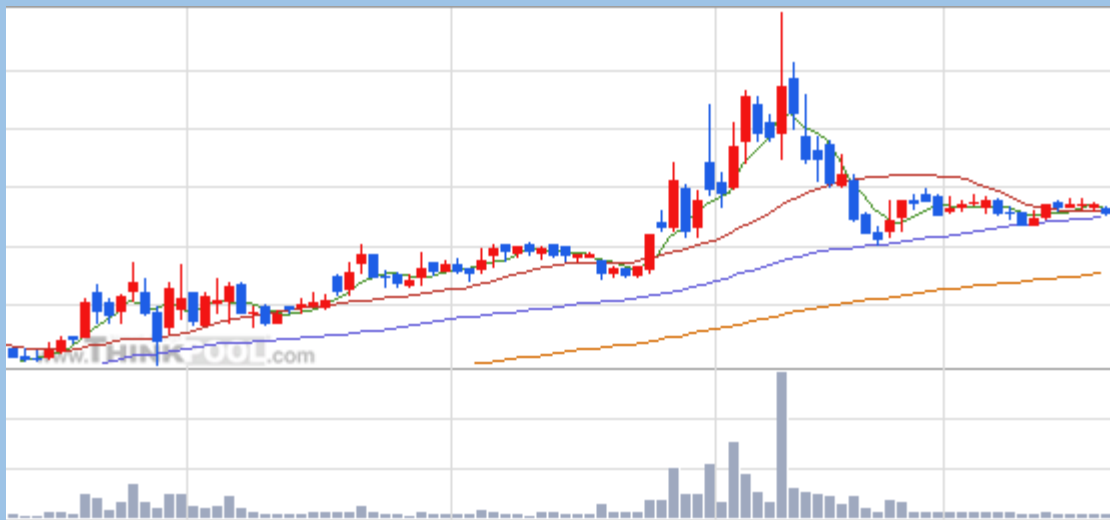


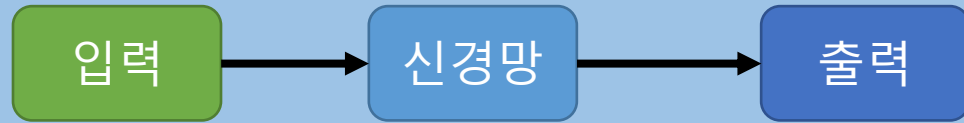
LSTM

시퀀스

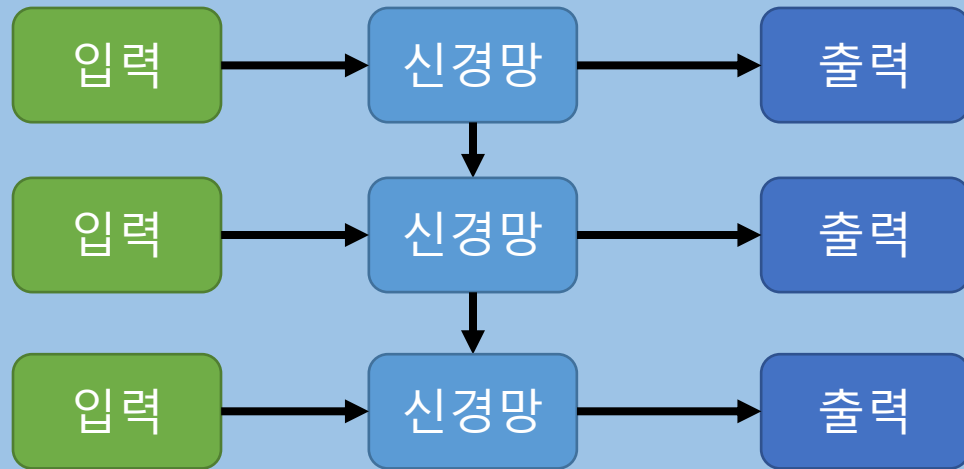


주식 차트와 같이
모델에 입력할 특징 변수가 순서에 따라
예측에 영향을 미치는 이른바 '시퀀스'
문제는 단일 벡터로 예측이 힘들다.

RNN



1번은 단일 가중치에 의해서 예측을 하는 형태이고,

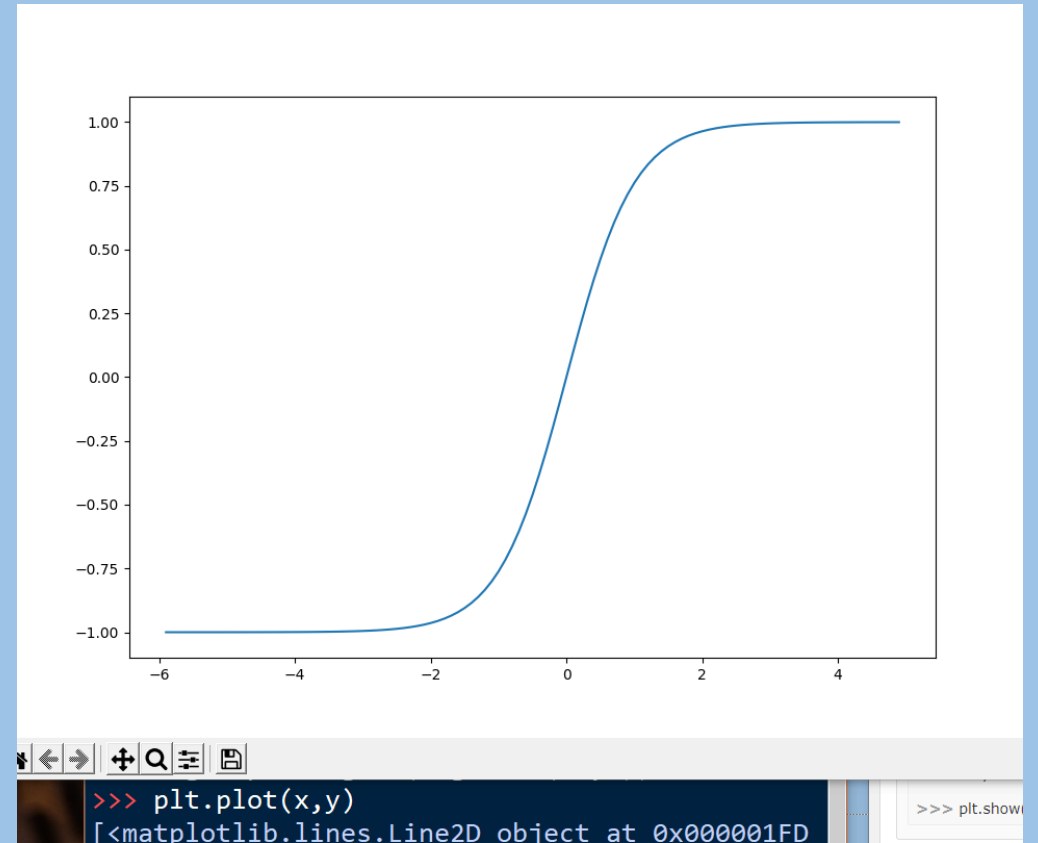


2번은 보는것과 같이 데이터의 순서대로 입력 받으며, 이전 출력이 다음 입력에 영향을 미치는 형태이다.

RNN

하이퍼 볼릭 탄젠트:

1. 시그모이드와 달리 $-1 \sim 1$ 의 값을 가진다.
2. 기울기 소실 문제가 조금 더 나아서 모델의 레이어를 쌓기 용이하다.
3. 또한 기울기의 범위가 양수, 음수 둘 다 가능하기 때문에 학습 효율성이 높다.



RNN 식

$$s_t = \tanh(s_{t-1} + x_t)$$

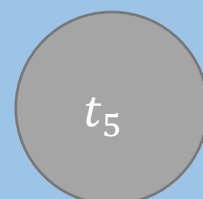
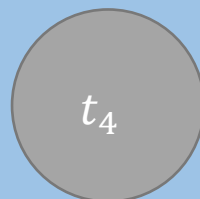
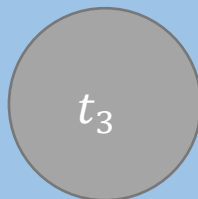
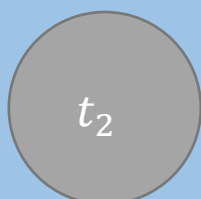
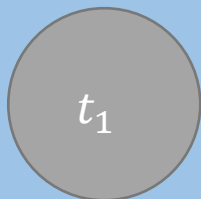
마지막 레이어

맨 마지막 레이어에는 0과 1사이의 확률값을 내놓는 시그모이드 함수를 사용한다.

단기 의존성 & 장기 의존성

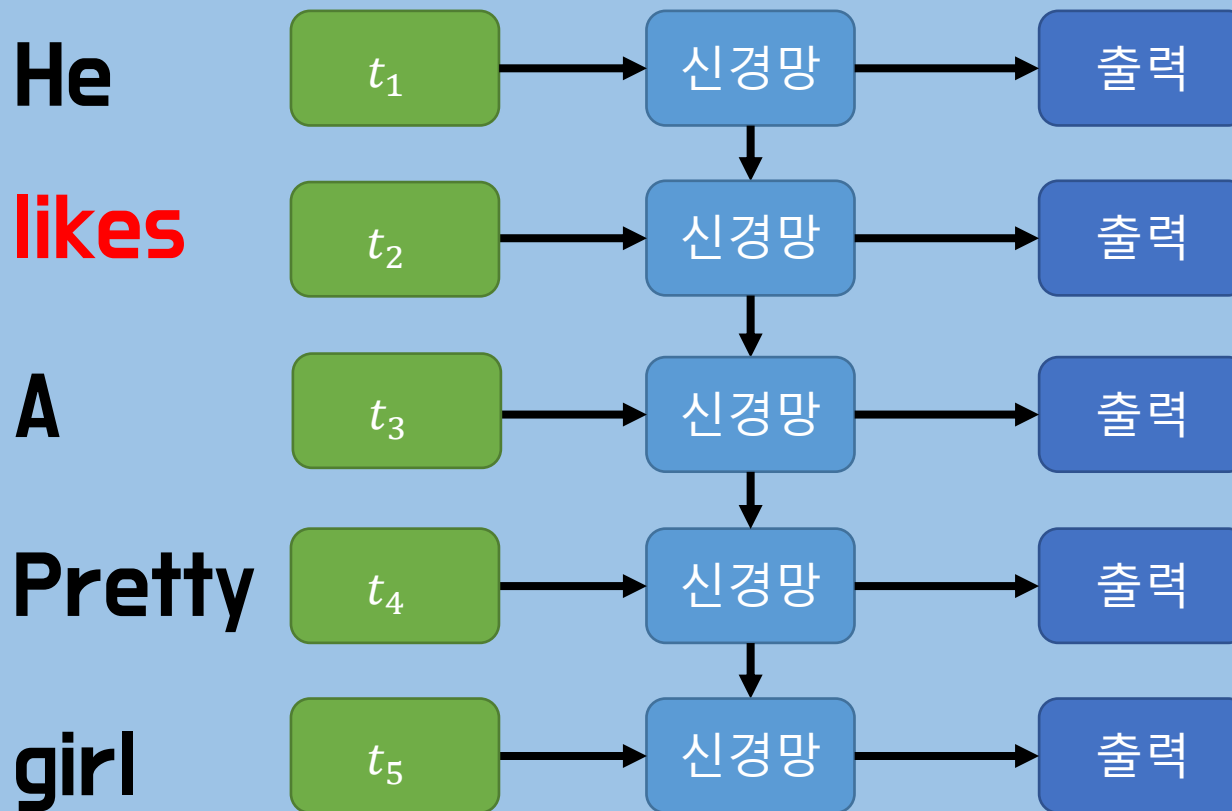
문장이 짧으면 주제를 잘 잡아내는데 문장이 길면 주제를 잡아 내지 못한다.
(= 단어의 개수가 많아질 수록 핵심 단어를 찾지 못한다.)

He likes a pretty girl



단기 의존성

이런 식으로 비교적 짧은 문장에서는
핵심단어를 파악하여 문장의 요지를
알아내는 것에 문제가 없다
-> 그가 여자를 좋아하는가? ○



장기 의존성

근데 만약에..

He seems to like that he likes a girl but he likes a boy
이러한 문장이 들어갔다고 생각해보자.

단어가 매우 많아지며 뜻을 해석하기 매우 난해해 할 것이다.

경사 소실 문제

역전파를 사용하려고 하여도, 오차값이 줄어들어 가중치에 명확한 영향을 주기가 어려워졌다.

LSTM

따라서 문장에서 주요한 부분만 얻어내어 요지를 판단하는 LSTM을 고안해 내었다.