

Recurrent Neural Network

강사 : 백병인

pi.paek@modulabs.co.kr

모두의연구소 Research Scientist



2019 모두의연구소

States in Sequence

Stateful 대화

손님: 안녕하세요.
점원: 어서 오세요. OO 햄버거입니다~.
손님: 햄버거 세트 하나 주세요.
점원: 사이드 메뉴는 무엇으로 하시겠습니까?
손님: 포테이토요.
점원: 음료수는 무엇으로 하시겠습니까?
손님: 콜라로 주세요.
점원: 50원 추가하시면 음료수를 L사이즈로 할 수 있는데 어떠신가요?
손님: 그냥 M사이즈로..
점원: 추가하실 건 없으시지요?
손님: 예
점원: 알겠습니다.

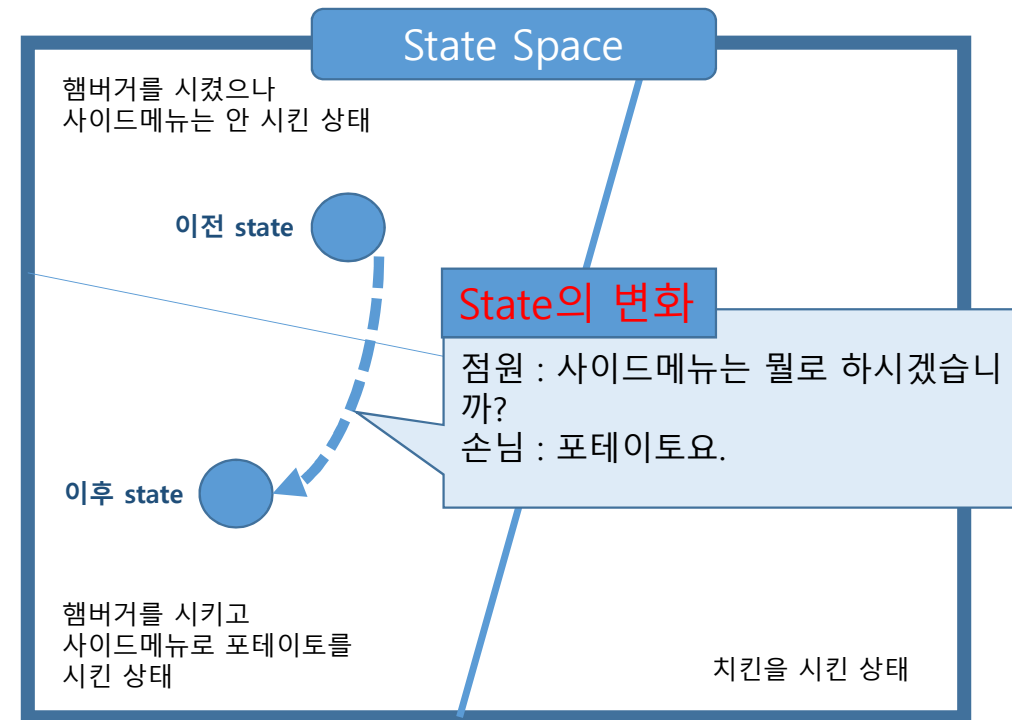
Stateless 대화

손님: 안녕하세요.
점원: 어서 오세요. OO 햄버거입니다~.
손님: 햄버거 세트 하나 주세요.
점원: 사이드 메뉴는 무엇으로 하시겠습니까?
손님: 햄버거 세트랑 포테이토 주세요.
점원: 음료수는 무엇으로 하시겠습니까?
손님: 햄버거 세트에 포테이토랑 콜라로 주세요.
점원: 50원 추가하시면 음료수를 L사이즈로 할 수 있는데 어떠신가요?
손님: 햄버거 세트에 포테이토랑 콜라는 M사이즈로..
점원: 추가하실 건 없으시지요?
손님: 햄버거 세트에 포테이토랑 콜라는 M사이즈로 부탁해요. 이상.
점원: 알겠습니다.

Stateless한 오른쪽 대화는 Sequence라고 볼 수 없다.
우리의 관심은 State의 변화를 담고 있는 Sequence를 모델링하고 싶은 것이다.

State란?

- state란? (<https://www.definitions.net/definition/state>)
 - **state, province(noun)**
 - the territory occupied by one of the constituent administrative districts of a nation
 - "his state is in the deep south"
 - **state(noun)**
 - the way something is with respect to its main attributes
 - "the current state of knowledge"; "his state of health"; "in a weak financial state"
 - **state(noun)**
 - the group of people comprising the government of a sovereign state
 - "the state has lowered its income tax"
- 종합하면, state란 어떤 공간(space)상에서 **경계선으로 구분된 영역**인데, 경계선 안쪽에서는 **동일한 지배자나 규칙에 의해 지배를 받는 영역**이다.
- 같은 경계선 안에서는 어디나 같은 영역이며, 같은 state로 규정된다. (같다와 다르다를 말할 수 있다는 뜻이다. 그리고 같다/다르다를 말할 수 있어야 변했다고 말할 수 있다.)

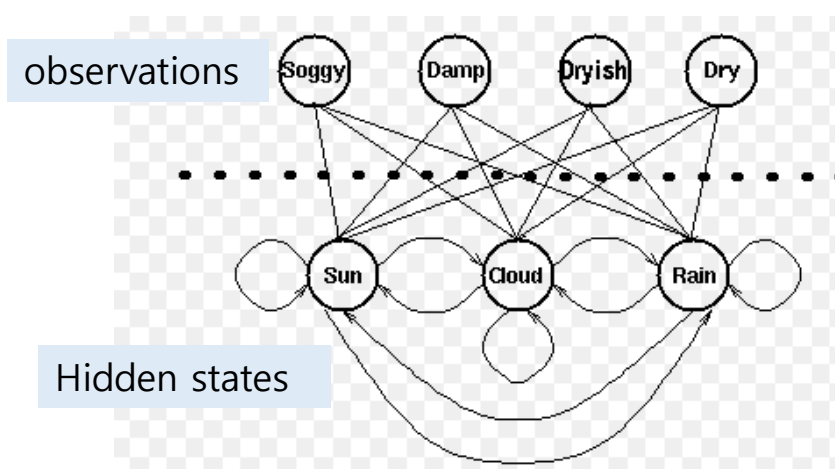


State의 변화로 발생한 점선 화살표를 순서대로 연결한 것을 우리는 Sequence라고 부른다.



Discrete or Continuous State

- State Space가 **유한집합**이라면?



State Diagram을 위 그림처럼 그릴 수 있을 것이다.
(State Space가 {Sun, Cloud, Rain}이라는 유한 집합이니까).

- State Space가 **무한집합**이라면?
 - 무지개의 색깔은 몇 개로 정의 가능할까?
 - 햄버거 가게의 state space는 몇 개로 정의 가능할까?
 - 대화하는 챗봇을 만들고 싶다. State space는 어떻게 정의해야 할까?
- State Representation : State를 **Vector**로 매핑하여 표현하자.
 - https://en.wikipedia.org/wiki/State-space_representation
 - State Representation이라는 것은, State는 아니지만 State를 표현하기 위해 도입한 Vector Representation이다.
 - Continuous State Space를 표현하기 위해서 이 Vector Space를 R^N 으로 둔다.

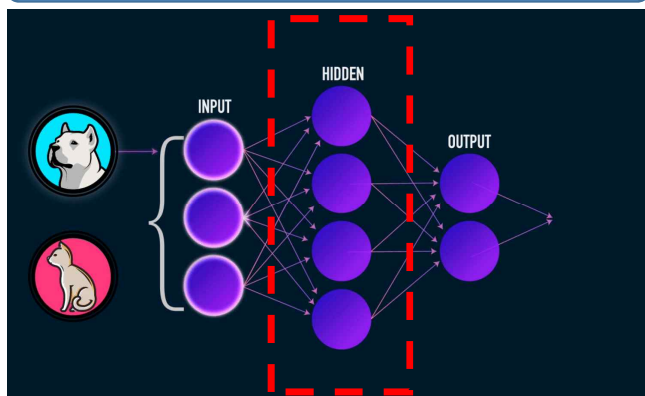
무한대의 state를 표현할 방법이 없어서, state를 N차원의 space에 맵핑하는 방법을 쓴다.
이 맵핑(Embedding or Encoding)이 정확/타당한지 항상 문제가 된다.



DNN : Continuous Mapping

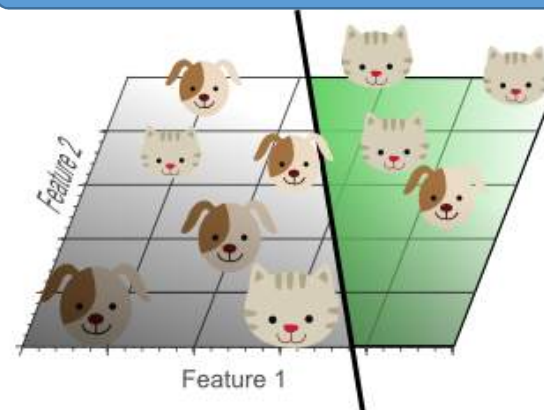
- 연결주의 모델인 Deep Neural Network는 Input을 Continuous (Hidden) Feature Space로 매핑하여, 그 space 상에서 경계선을 찾는 방식을 (통계적으로) 구현한다.

입력을 다른 연속공간으로 매핑



<https://www.analyticsindiamag.com/how-to-create-your-first-artificial-neural-network-in-python/>

연속공간 위에서 경계선 찾기

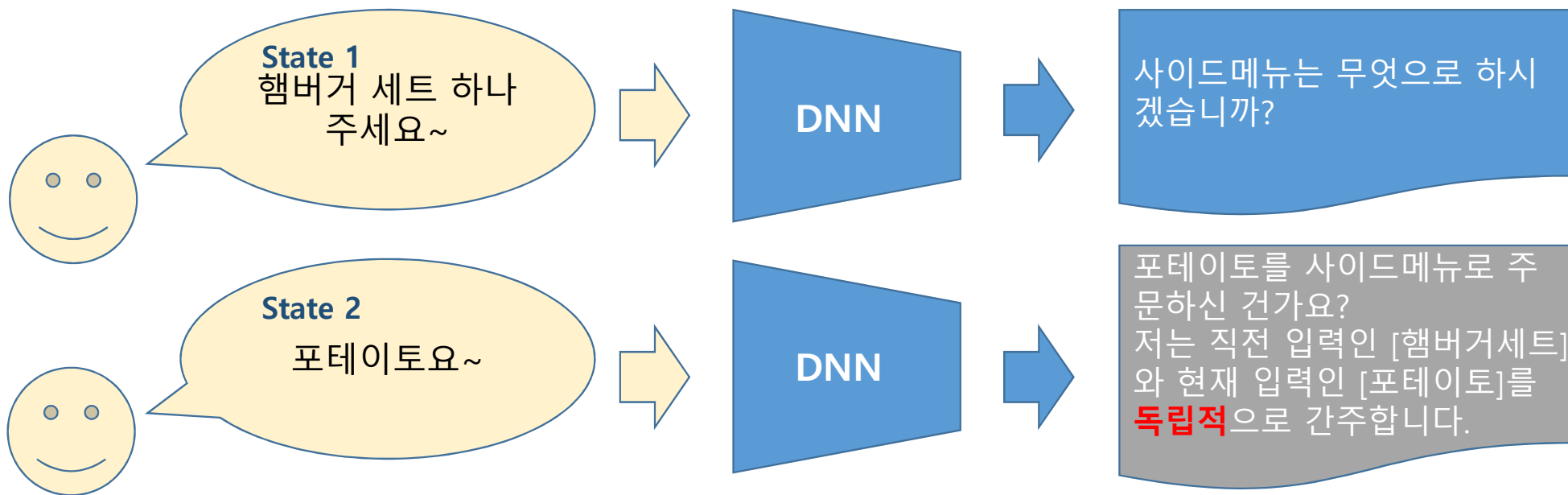


<https://www.cnblogs.com/oon/p/4492168.html>

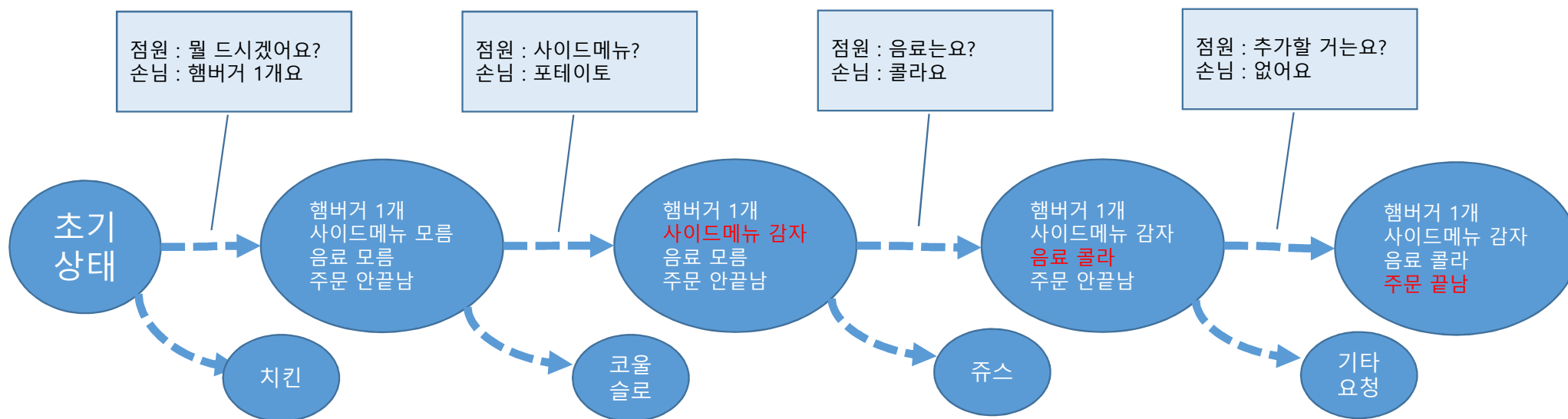
- 귀납적 통계추론의 대원칙 : 그러나, Backpropagation을 통한 Hidden Space 매핑이 바르게 이루어지려면 반드시 입력데이터 (개사진, 고양이사진...)들끼리는 서로 **독립적**이어야 한다.

DNN으로는 Stateful Machine을 구현할 수 없다.

Deep Neural Network 점원



Discrete State Machine



HMM(Hidden Markov Model)으로 점원을 구현하려면, 모든 state를 다 정의해 주어야 한다.
몇 개의 state를 정의해야 충분할까?

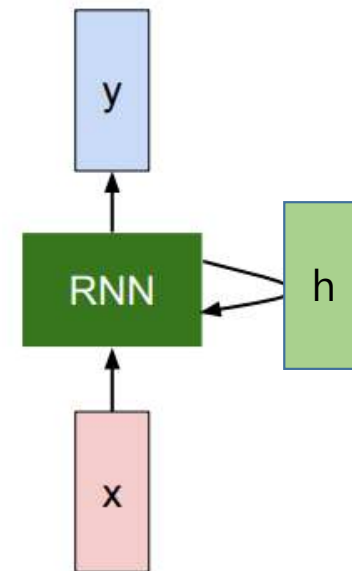
RNN은 Continuous State Machine이다.

- State를 (discrete한) 동그라미가 아니라 벡터로 표현했다.
- 하지만 표현하고 싶은 것은 입력 x 의 변화에 의한 State h 의 변화이다.

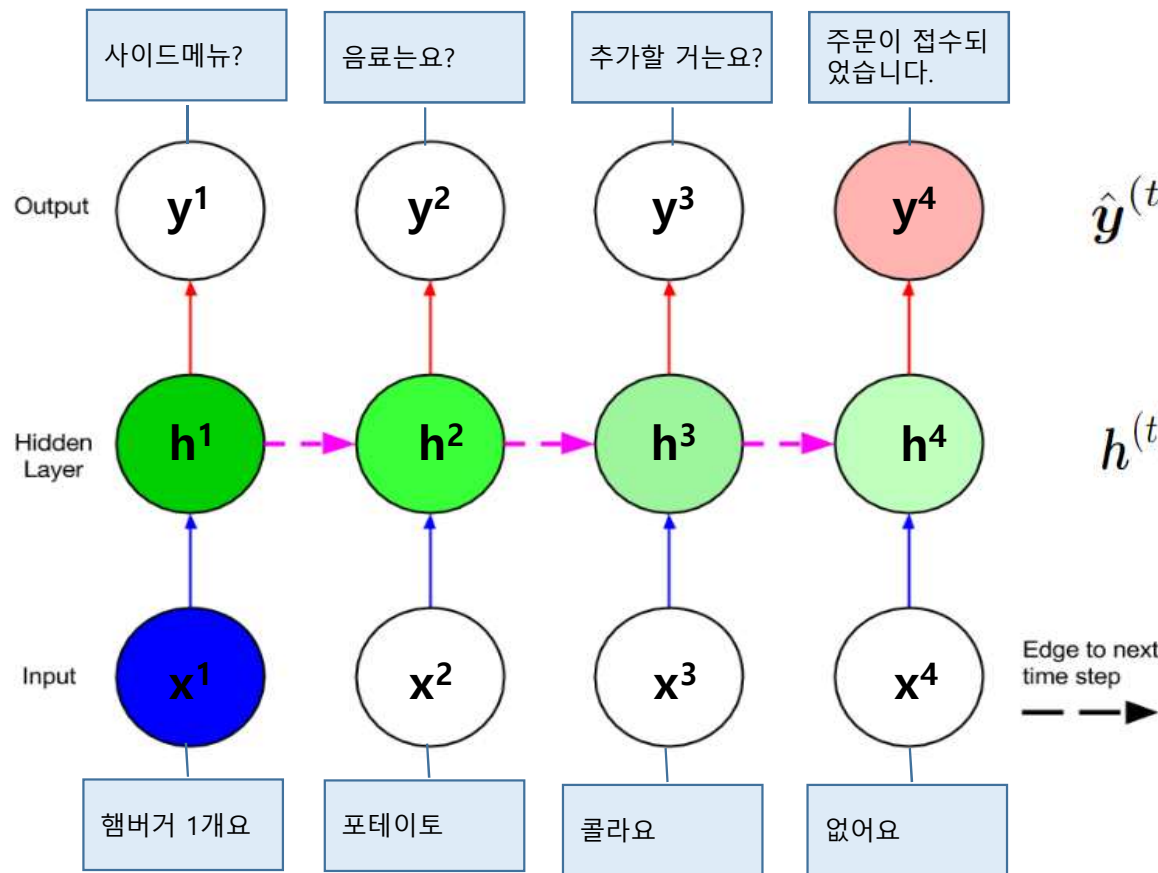
We can process a sequence of vectors x by applying a **recurrence formula** at every time step:

$$\boxed{h_t} = \boxed{f_W}(\boxed{h_{t-1}}, \boxed{x_t})$$

new state some function with parameters W old state input vector at some time step



RNN으로 구현한 Stateful 점원



출력은 **현재상태**가 결정합니다.

$$\hat{y}^{(t)} = \text{softmax}(W^{yh}h^{(t)} + b_y)$$

RNN의 출력(점원의 대답) y^t 은 새로운 상태벡터 h^t 에 가장 적절한 것으로 선택된다.

$$h^{(t)} = \sigma(W^{hx}x^{(t)} + W^{hh}h^{(t-1)} + b_h)$$

새로운 상태벡터 h^t 는 이전 상태벡터 h^{t-1} 과 새로운 입력벡터 x^t 의 벡터 선형결합(linear combination)으로 결정된다.

현재상태는 **입력**과 **직전상태**가 결정합니다.

RNN cell & network

