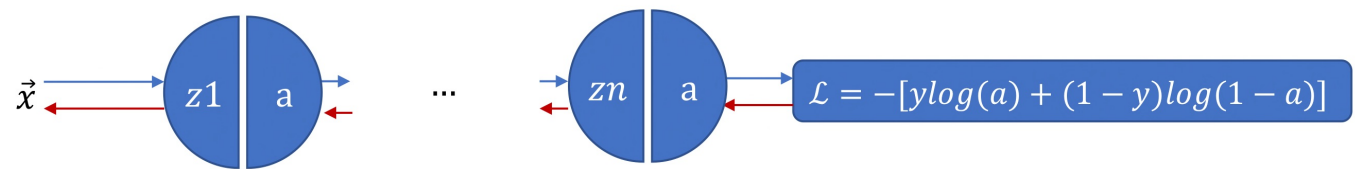


### Question. 7-07

다음과 같이 ReLU Function을 이용하여 Logistic Regression Multi-Layer를 쌓았다고 가정하자.

Sigmoid Function, Tanh Function으로 쌓은 경우와 비교하여 개선된 점과 문제점을 찾고 그 이유를 설명하라.



1) Backpropagation 진행 시 ReLU Function이 의해 updates되는 값이 작아지지 않는다.

따라서 Vanishing gradient problem이 해결되고 Sigmoid, Tanh Function이 비해 빠른 연산 속도를 가진다.

그러나 출력 0에서 미분 불가능한 함수로 인해 0에서 Backpropagation 시 기울기를 출력할 수 없다.

하지만 컴퓨터가 정확히 0으로 가는 경우가 드물기 때문에 경로를 나눠 미분값을 사용한다.

또한 편미분값이 0이면 해당 층이 updates가 되고 국소최대라는 경우가 생기거나 크게 중요해지지 않는다.