

그럼 Embedding 레이어는?

Embedding 레이어는 ^①처음에 무작위로 초기화된 상태에서 정수로 오는 word를 정해진 크기의 벡터로 바꿔서 다음 레이어로 넘기고, ^②~~학습~~^③단계에서는 역전파되는 기울기를 바탕으로 해당 word의 임베딩 값을 조정한다. 즉 주변 문맥을 반영하지 않는다.

↑ word2vec 다름 다른점.

값을 벡터로 바꾸는 과정이, 가중치 행렬이 존재한다.

그럼 Dense레이어랑 사실상 똑같은거 아닌가? 싶은데 차이점은 원핫벡터로 값을 안넣어줘도 되서 덕분에 메모리 절약할 수 있고, 케라스에서 지원하는 masking 기능을 사용할 수 있고, 꼭 문장뿐만 아니라 추천 시스템 등에서 User 등을 벡터로 나타낼 때도 사용할 수 있다.

하지만 어지간하면 weights 지정 기능을 이용해서 pre-trained 된 word2vec 등을 지정해야 할 것이다.