2019150445 통계학과 신백록

1.

LSGAN 모형은 다른 GAN 모형과 다르게 loss function이 binary crossentropy에 기반을 두고 있지 않다. Loss가 binary crossentropy에 의해 도출된다면 3번 문제처럼 data의 분포와 generator의 분포가 동떨어져 있을 때 loss가 상수 값이 나와 더 이상 update가 되지 않는다는 문제가 있다. 하지만 LSGAN의 discriminator의 loss는 로, 만 아니면 loss는 항상 update가 된다. 이는 generator의 loss function에 영향을 끼치고, generator를 개선한다.

2.

LeakyReLU함수는 ReLU와 다르게, 값이 0보다 작을 때 0을 출력하는 것이 아니라 음수 값을 출력한다. WGAN모형에서는 discriminator와 generator의 출력 값이 -1에서 1사이이기에 ReLU함수를 사용하여 음수 값을 0으로 보내면 안 된다. WGAN모형에서 ReLU함수만을 사용한다면 출력층에서 tanh activation을 사용할 때 0에서 1사이 값만 나와 항상 1(real)이 나올 것이기 때문이다.

또한 ReLU함수는 dropout의 효과도 있기에 WGAN모형에서 ReLU함수를 사용하지 않는 대신 모수의 숫자가 적음에도 dropout을 넣어준 것 같다.

3.

에서 일 때, 즉 x1이 오직 1의 값만을 가질 때 이 을 따른다고 가정하자.

마찬가지로, 에서 일 때, 즉 x1이 오직 2의 값만을 가질 때 이 을 따른다고 가정하자.

따라서 .와 는 x1이 가지는 값이 다르기에 겹치는 부분이 아예 없다. 따라서 한 분포가 U(0,1)을 따른다면(1의 값을 갖는다면) 다른 분포는 0의 값을 갖는다.

에서 는 x1=2 하에서 U(0,1)을 따르기에 에서의 적분 범위는 0부터 1까지고, 는 0 는 1로 고정된다. 을 0부터 1까지 적분하면 무한대가 나오고, 도 마찬가지로 계산된다.

도 마찬가지로, 0부터 1까지 의 값, 의 값을 갖는다. 따라서 의 값이 나오고, 도 마찬가지로 의 값이 나온다.

4.

BatchNormalization은 학습을 시행할 때 각 layer의 노드들의 평균과 분산을 가지는 정규분포로 normalization을 시행한다. 하지만 test set에 대하여 이를 시행한다면 model을 학습할 때의 BatchNormalization을 통해 분포를 추정하는 것의 의미가 없어지게 된다. 따라서 모델을 학습할 때 mini batch의 이동평균과 분산을 저장하고, test set에 대하여 각 mini batch에서 구한 평균과 분산을 사용해 momentum에 따라 이동평균을 계산하게 된다.