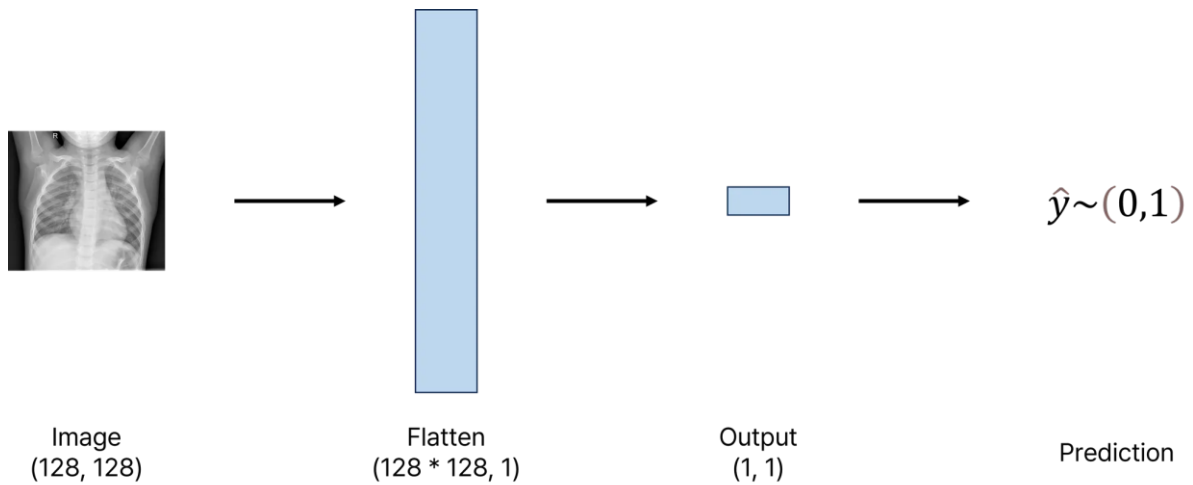


1. 모델 구조 요약

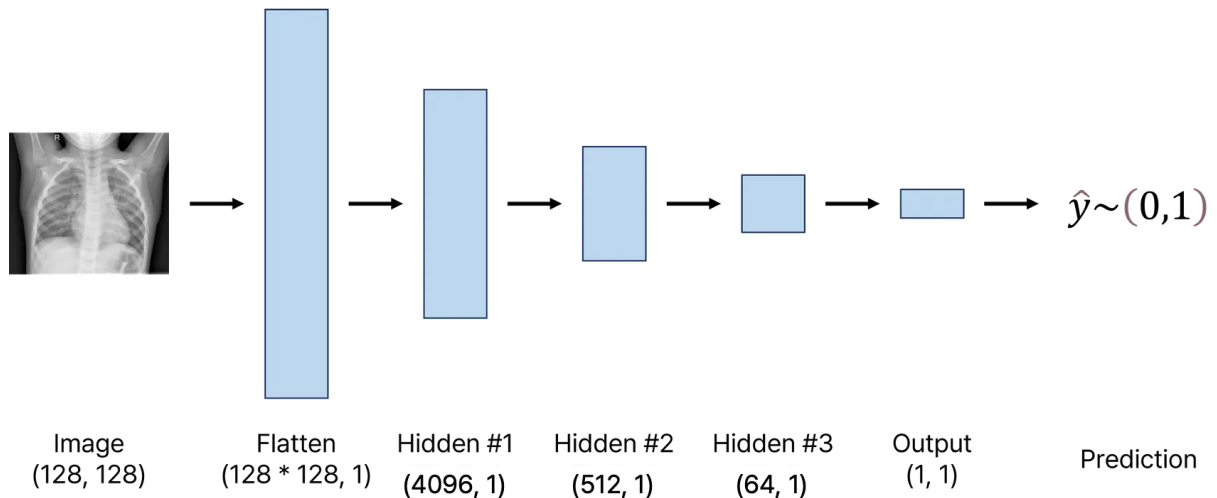
A. Logistic Regression (LogReg)

Logistic Regression 모델은 linear layer를 이용해 $128 * 128$ 개의 pixel을 1개의 값으로 바꿔주고, 이를 sigmoid 함수를 통해 0과 1 사이의 값으로 mapping하였다.



B. Deep Neural Network (DeepNN)

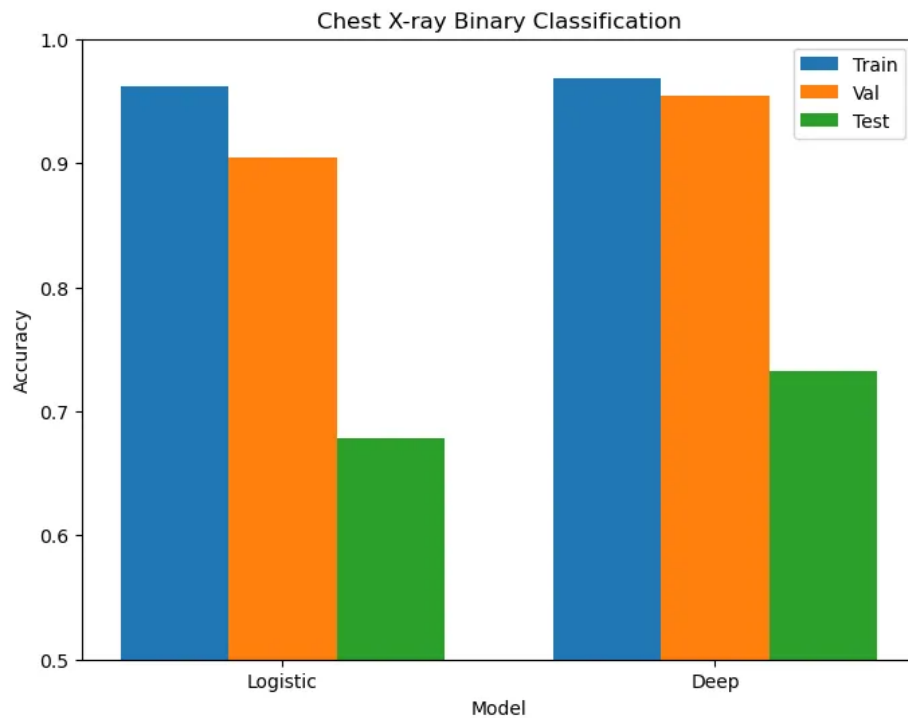
Deep Neural Network(DeepNN) 모델은 input dimension을 $4096 \rightarrow 512 \rightarrow 64 \rightarrow 1$ 의 순서로 linear multiplication을 수행하였다. activation function은 hidden layer에서는 모두 ReLU 함수를 사용하였고, output layer에서는 Binary classification을 위해 sigmoid 함수를 사용하였다.



2. 정확도 비교 및 해석

A. Bar Plot을 통한 정확도 비교

Epoch = 20일 때의 정확도를 비교한 결과 다음의 Bar plot을 얻을 수 있었다.



B. 정확도 차이의 원인 분석

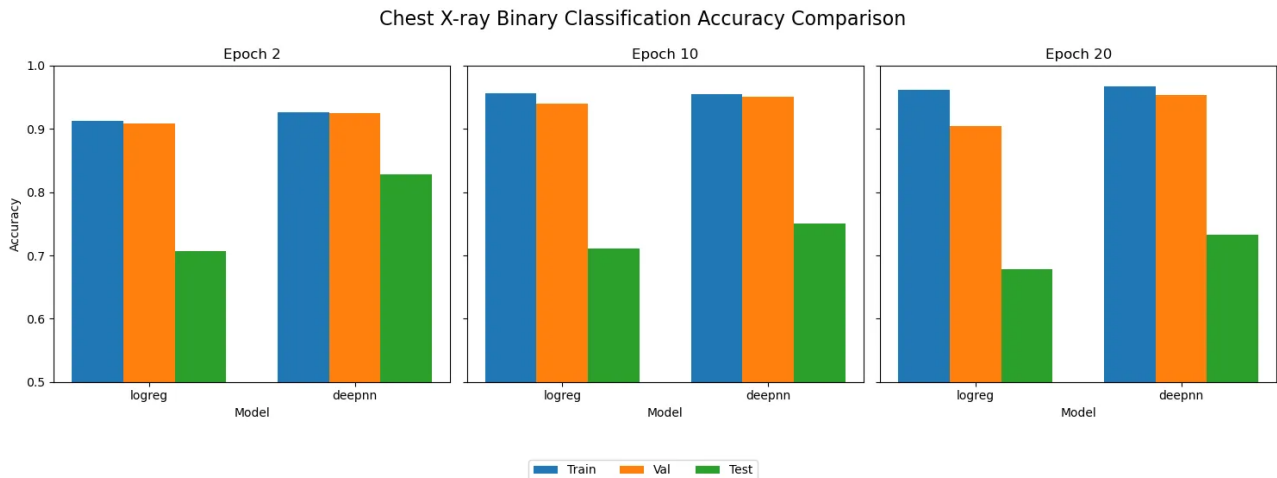
LogReg은 단일 linear 계층과 sigmoid 함수로만 이루어진 반면, DeepNN은 ReLU 계층으로 인해 비선형 변환이 가능하다. 즉, 복잡한 패턴을 학습하는 데 있어 DeepNN이 유리하다.

또한 Parameter의 수를 비교했을 때 Logistic model은 $128 * 128 = 16,384$ 개가 전부이지만, DeepNN은 $(16,384 * 4,096) + (4,096 * 512) + (512 * 64) + (64 * 1)$ 개로, 학습할 수 있는 양이 더 많기 때문에 DeepNN이 더 높은 정확도를 보일 수 있다.

3. 추가 분석

A. Epoch = 2, 10, 20일 때의 정확도

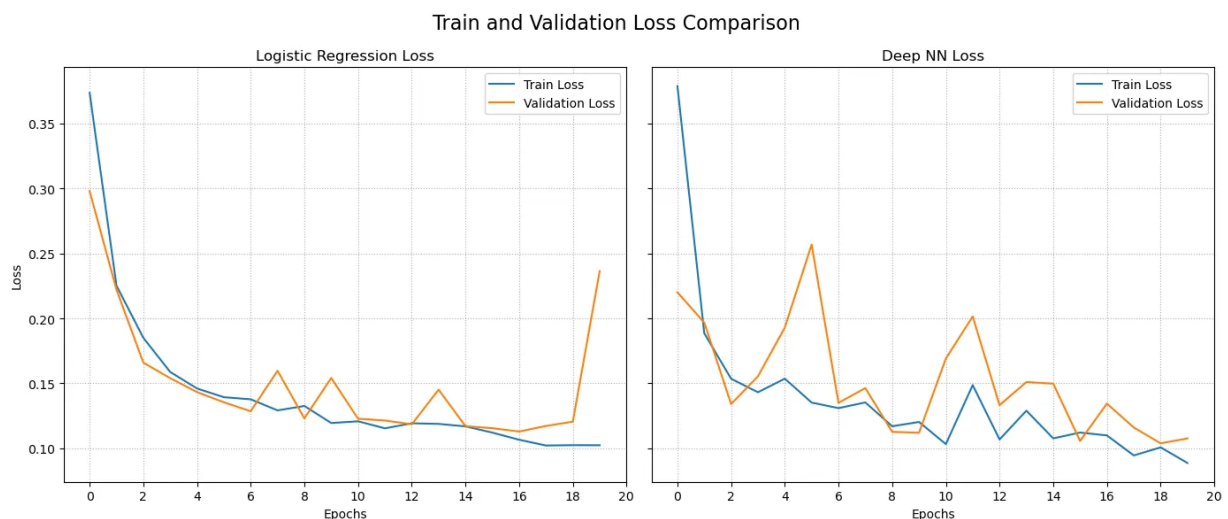
분석을 원하는 epoch에서 test를 실행하도록 코드를 작성했고, 그 결과 다음의 bar plot을 얻을 수 있었다.



Train과 validation set에 대해서는 두 모델 모두 점진적으로 증가하는 경향을 보였으나, LogReg의 validation accuracy가 낮아진 것으로 보아 두 시점 사이에 overfitting이 발생했음을 추론할 수 있다. test set에서의 정확도가 epoch = 20에서 낮아지는 점도 overfitting에 따른 결과로 해석된다. 다만, DeepNN의 정확도가 낮아지는 경향성에 대해서는 뒤의 Loss 함수에 대한 그래프를 보며 같이 다루겠다.

B. 각 Epoch에서의 overfitting, underfitting

Epoch마다의 train loss, validation loss를 측정하여 어느 시점에서 overfitting이 발생했는지를 확인하고자 한다. 아래 그래프는 각 epoch에서의 loss를 측정한 그래프이다.



LogReg의 경우 validation loss가 꾸준히 감소하다가 epoch = 19에서 20으로 진행될 때 loss가 크게 증가한 것으로 보아 epoch = 19 이후에 overfitting이 일어났음을 진작할 수 있다.

반면, DeepNN의 경우 validation loss의 변동이 크지만 꾸준히 감소하는 경향성을 보인다. 변동이 큰 이유는 모델이 Logistic regression에 비해 더 복잡하기 때문으로 고려되며, validation loss의 경향성과 앞의 bar plot으로 미루어 보았을 때 underfitting으로, 더 많은 epoch나 더 큰 learning rate가 필요하다고 해석된다.