**HW#2 – Neural Network Experiment**

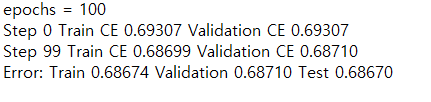
**2018/06/11**

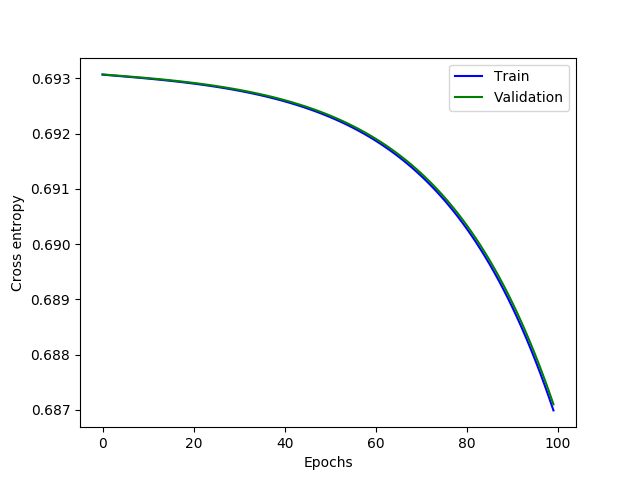
**CSE 201624476 박상운**

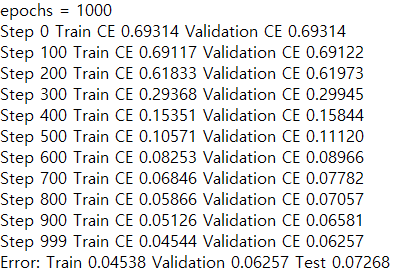
1. **Basic Generalization**

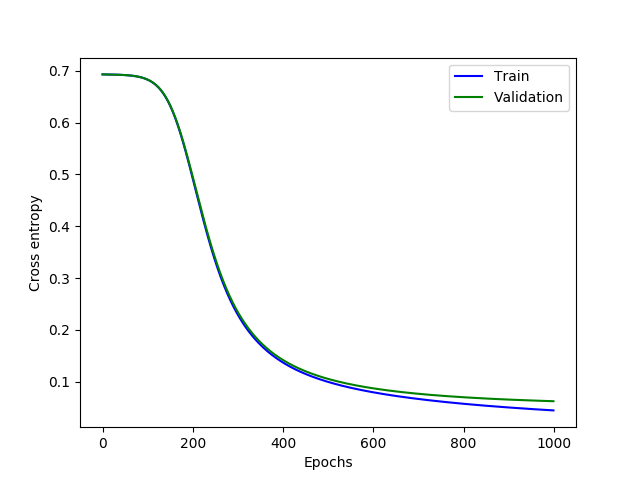
Epoch value에 따른 학습 경향의 변화를 각각 N= { 100, 1000, 10000 } 일 때에 대해서 시행해보았습니다.

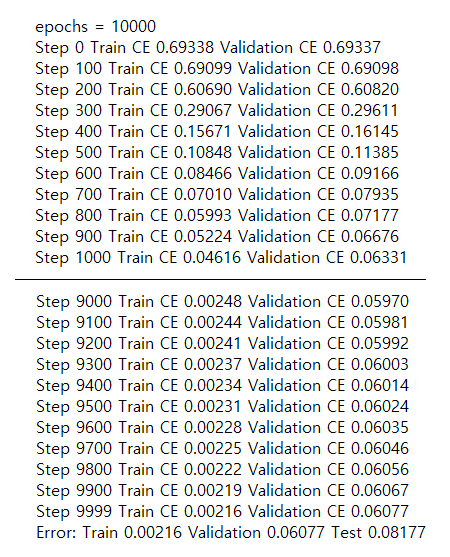
(Epoch value를 제외한 모든 조건은 eps = 0.1, num\_hiddens = 10, momentum = 0.0, class entropy measured으로 동일합니다.)

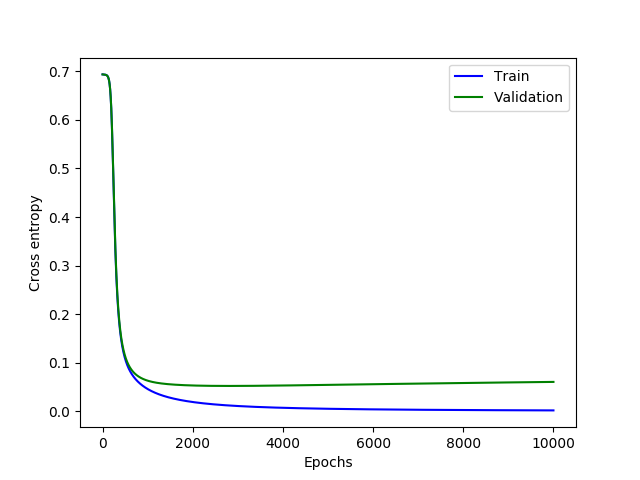










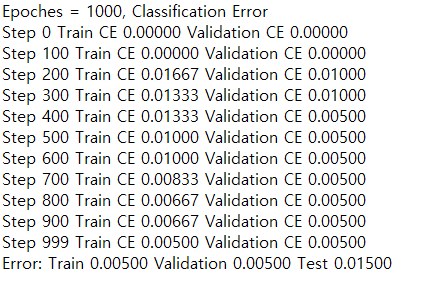


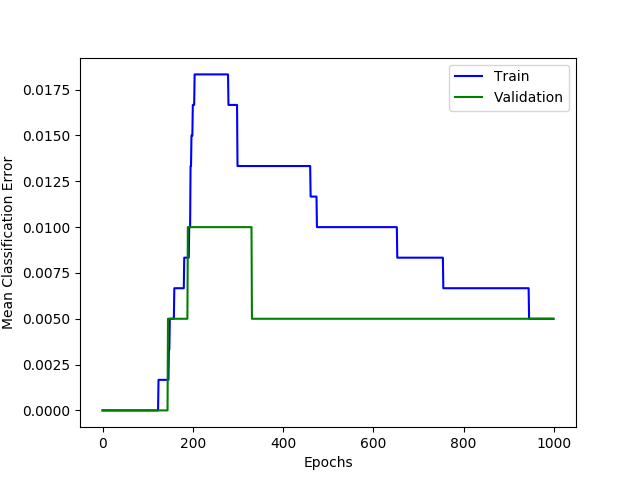
N = 10000 model을 통해서 Epoch value가 대략 1000을 넘어가는 부분을 기점으로 Train error는 감소하지만 validation error는 오히려 증가하는 모습을 볼 수 있습니다. 이를 근거로 epoch value가 과하게 설정될 경우 Overfitting이 발생함을 판단할 수 있습니다.

1. **Classification Error**

classification error는 각각의 prediction에 대해서 True/False값을 가집니다. 예를 들어 Prediction vector가 (0.3 0.4 0.3) 일 때, label vector가 (0, 1, 0)으로 맞으면 False 값을 갖고, (1, 0, 0)으로 틀리면 True 값을 갖습니다. 각각이 0,1의 값을 가질 때 이에 대한 평균을 낸 것이 Mean classification error입니다. 예를 들어 Prediction 8개 중 3개가 맞을 경우 5개의 True와 3개의 False가 발생하며 이에 대한 평균은 0.625입니다.

Cross Entropy를 Mean Classification Error로 대신하여 실행한 결과는 다음과 같습니다.

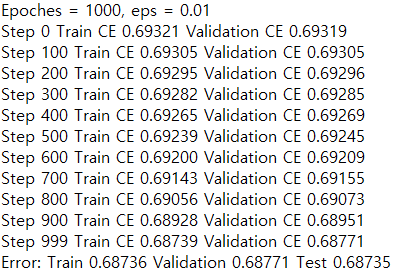


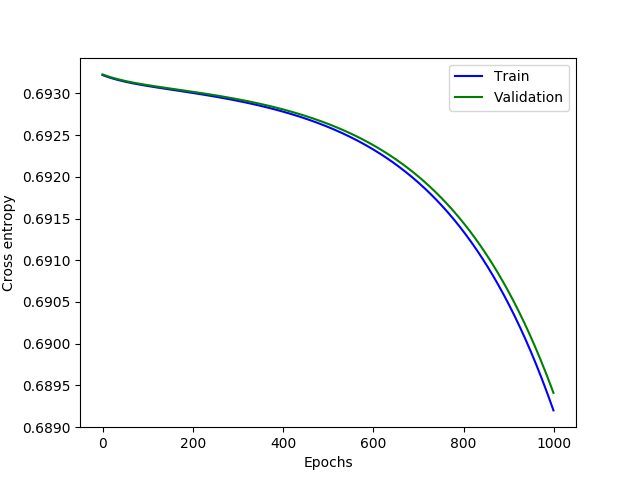


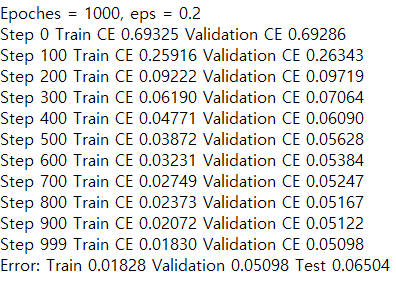
그래프에 보이는것과 같이 Mean Classification Error를 사용할 경우에는 그래프가 Discrete한 분포를 보이며, 이때 epoch의 개수만큼 0~1의 구간이 나눠 짐을 확인할 수 있습니다. Classification error는 정답률에 대한 정보만 제공하며 정답의 정확도와 goal state에 대한 distance 값을 제공하지 않으므로 탐색 과정에 있어서 신용도가 낮으므로 위와 같이 error rate가 0에서 급증하게 되는 결과가 생깁니다. 따라서 평가 함수로 부적합하다고 판단할 수 있습니다.

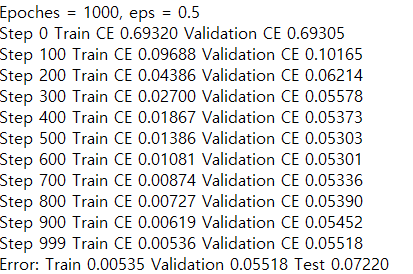
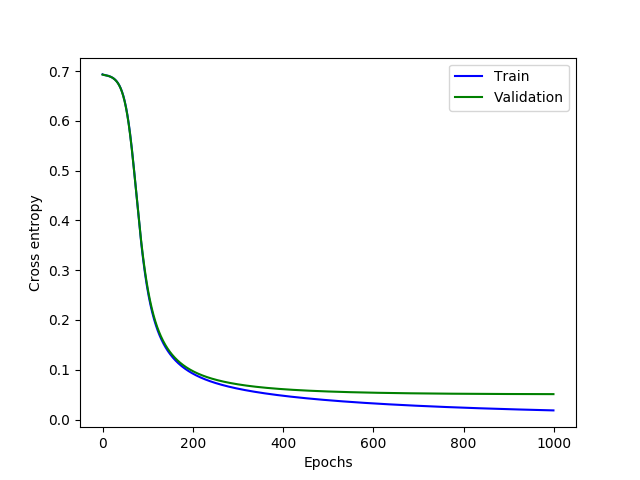
1. **Learning rate**

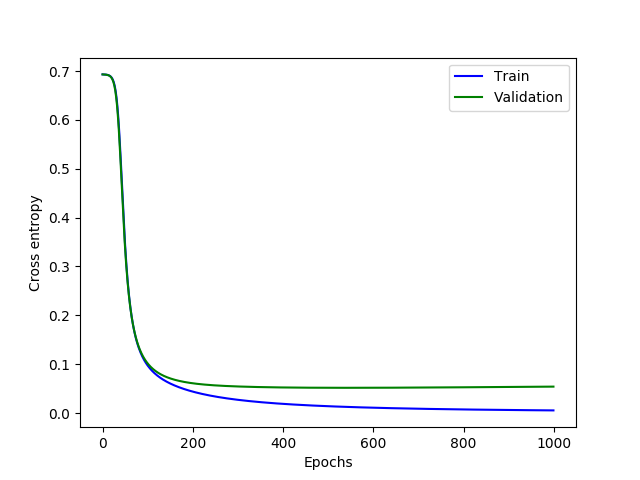
Eps value는 기본적으로 0.1로 설정되어 있으며, 이를 {0.01, 0.2, 0.5}로 설정한 후 epoch=1000, num\_hiddens = 10, momentum = 0에서 각각 시행한 결과는 다음과 같습니다.





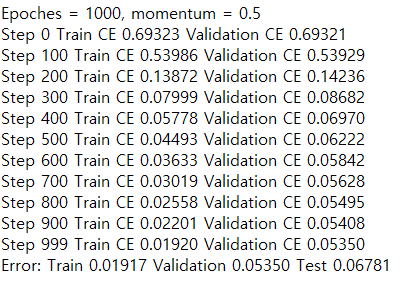


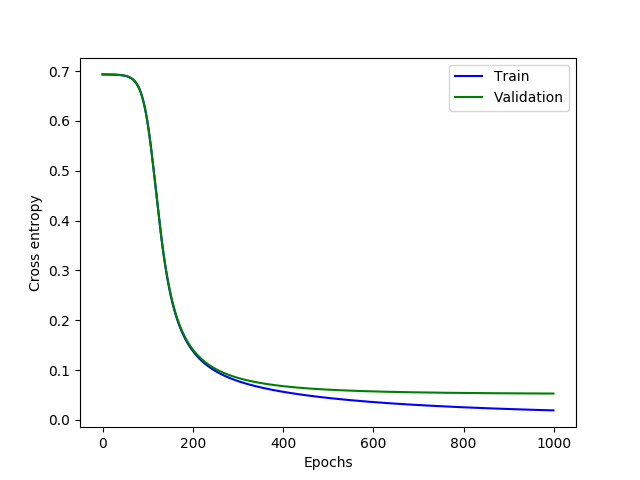


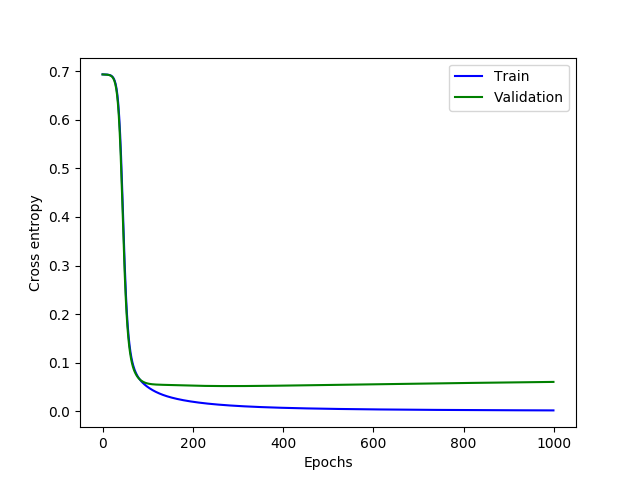
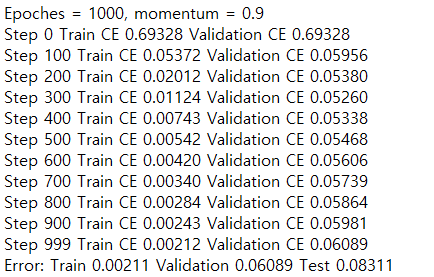


Eps의 증가는 Epoch value의 증가보단 덜하지만 Training data set에 대해서 최적화되는 경향을 살펴볼 수 있으므로 값을 과하게 증가시킬 때에 대한 overfitting의 가능성을 내포하고 있다고 판단할 수 있습니다.

Momentum value의 경우에는 기본값 0.0이 아닌 {0.5, 0.9} 두 값에 대해 각각 시행하였으며 이 외의 조건들은 epoch=1000, eps=0.1, num\_hiddens=10으로 동일합니다.



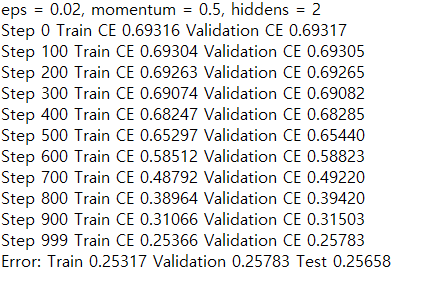


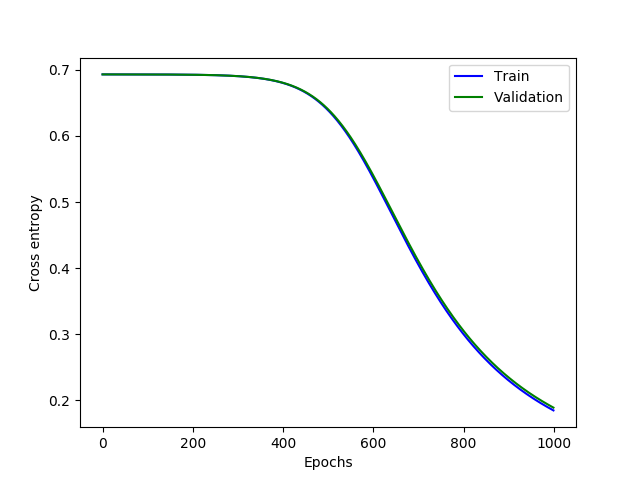


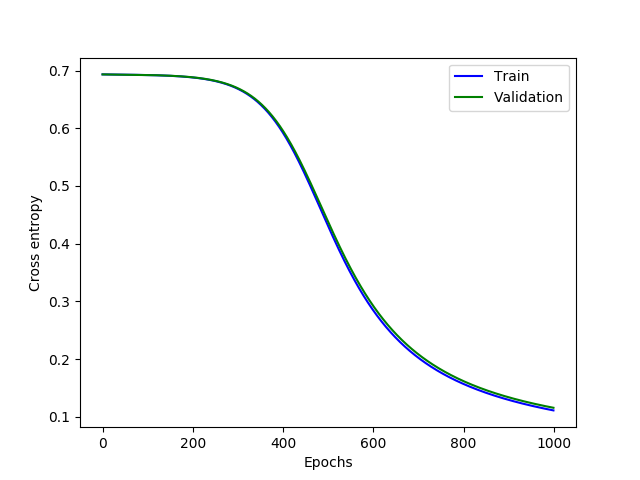
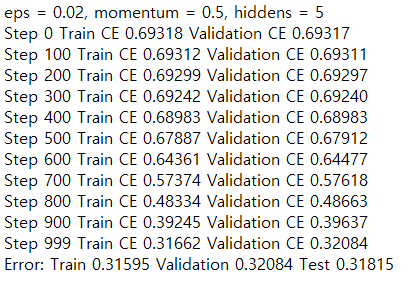
이에 따라 momentum이 증가할수록 error graph의 기울기 변화량이 가파르게 변함을 알 수 있습니다. 또한 충분히 큰 동일 epoch value에서 overfitting이 심해짐을 확인할 수 있습니다.

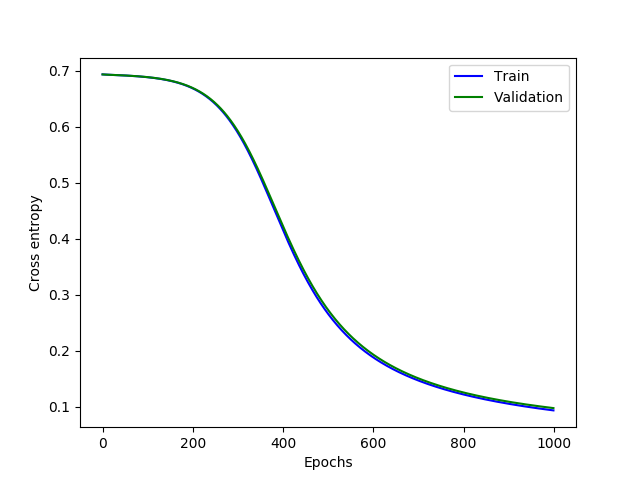
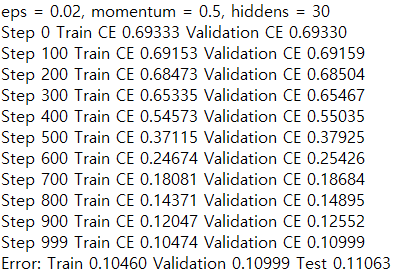
1. **Number of hidden units**

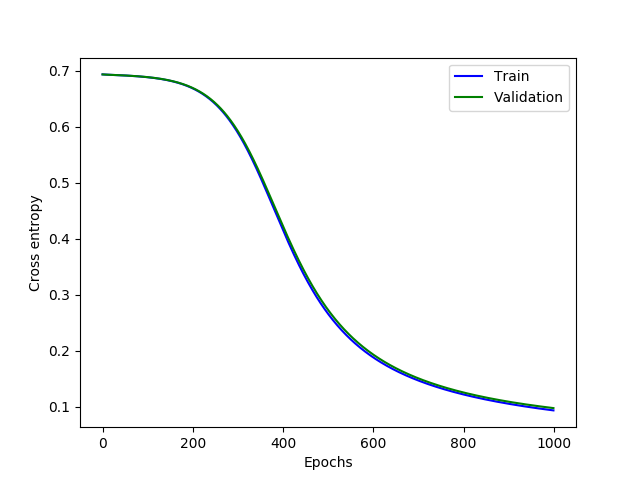
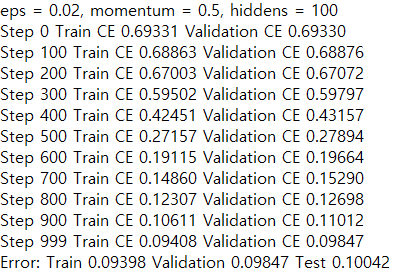
Eps=0.02, momentum = 0.5 상태에서 hidden unit 개수를 {2, 5, 30, 100} 으로 설정한 결과는 다음과 같습니다.











다른 요인들에 비해서 값이 증가하여도 overfitting이 발생하는 정도가 덜한 경향을 볼 수 있습니다. 또한 validation set의 cross entropy 또한 확실하게 감소하는 경향을 확인할 수 있습니다.