**DL\_HW1\_정주현\_21512023**

1. **데이터셋 이해**

Fashion MNIST 데이터셋은 운동화, 셔츠, 샌들과 같은 흑백 이미지들의 모음이며, 10가지로 분류될 수 있는 28\*28 픽셀의 이미지 70,000개(train : 60,000개, test : 10,000개)로 이루어져 있다. 클래스 정보: (0: T-shirt/Top, 1: Trouser, 2: Pullover, 3: Dress, 4: Coat, 5: Sandal, 6: Shirt, 7: Sneaker, 8: Bag, 9: Ankle Boot)

1. **4가지 모델 설명**
   1. Logistic Regression

입력 28\*28, 출력 10인 1개의 linear layer로 구성되어 있다.

* 1. Feedforward Neural Network1 (FNN1)

입력층(노드 28\*28개), 은닉층1(256개), 은닉층2(64개), 출력층(10개)로 구성되어 있다. 입력층,->은닉층1 계산 시 활성화함수 ReLU를 적용하고, 은닉층1->은닉층2 계산 시 ReLU를 적용하고 20%비율로 Dropout을 적용한다.

* 1. Feedforward Neural Network2 (FNN2)

학습하는 과정 자체를 전체적으로 안정화하여 학습 속도를 가속시킬 수 있으며 gradient vanishing, exploding문제를 방지하는 Batch Normalization를 이용하였다.

입력층(노드 28\*28개), 은닉층1(256개), 은닉층2(64개), 출력층(10개)로 구성되어 있다. 입력층,->은닉층1 계산 시 Batch Normalization을 수행하고 활성화함수 ReLU를 적용한다. 은닉층1->은닉층2 계산 시 Batch Normalization을 수행하고 ReLU를 적용한다.

* 1. Feedforward Neural Network3 (FNN3)

입력층(노드 28\*28개), 은닉층1(256개), 은닉층2(64개), 출력층(10개)로 구성되어 있다. 입력층,->은닉층1 계산 시 Batch Normalization을 수행하고 활성화함수 ReLU를 적용한다. 은닉층1->은닉층2 계산 시 Batch Normalization을 수행하고 ReLU를 적용하고 20%비율로 Dropout을 적용한다..

1. **세팅한 hyperparameter의 의미**

batch size = 64 : 한번의 batch(나눠진 데이터셋)마다 주는 데이터 sample size

lr = 0.01 : gradient의 학습률

epoch = 10 : 전체 데이터를 10번 사용해서 학습을 거친다.

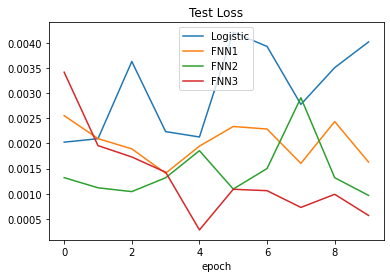
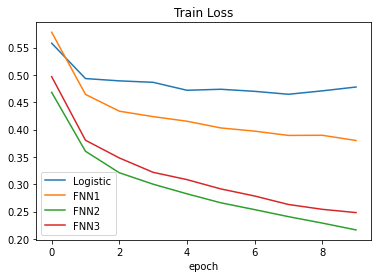
즉, 한 epoch당 data개수/64 번의 iteration을 거친다.

1. **활용한 2가지 regularization과 그것에 대한 설명**

Dropout : overfitting을 방지하기 위해 학습할 때 서로 연결된 layer에서 0부터 1 사이의 확률로 뉴런을 제거하는 기법

Batch Normalization : Batch 단위로 학습을 하게 되면 Batch 단위 간에 데이터 분포의 차이가 발생할 수 있으므로 각 배치별로 평균과 분산을 이용해 정규화한다. 이를 통해 학습하는 과정 자체를 전체적으로 안정화하여 학습 속도를 가속시킬 수 있으며 gradient vanishing, exploding문제를 방지한다.

1. **epoch이 진행됨에 따른 train loss, test loss의 변화 그래프 포함**

****

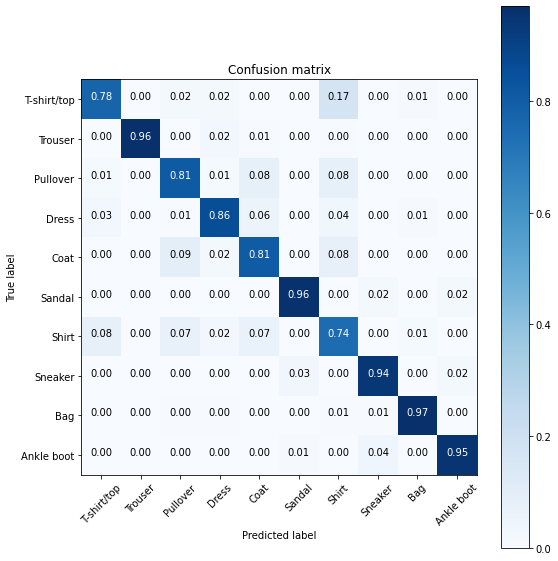
1. **각 모델의 test set에 대한 성능 확인 및 가장 우수한 모델 선정**

텍스트, 클립아트이(가) 표시된 사진

자동 생성된 설명

가장 우수한 모델은 FNN3이다.

1. **해당 모델에 대한 confusion matrix 확인 및 가장 분류가 어려운 class는 무엇인지 확인**



FNN3에 대해 normalized confusion matrix을 출력한 결과 가장 분류가 어려운 class는 Shirt이다. Shirt을 T-shirt/top으로 오분류하는 경우가 많았다.

**코드 첨부**

텍스트이(가) 표시된 사진

자동 생성된 설명 -텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명 텍스트이(가) 표시된 사진

자동 생성된 설명 텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명 텍스트, 테이블이(가) 표시된 사진

자동 생성된 설명 