

과제 #6

M3239.002300_003 데이터 사이언스 특강

Due: 2021년 11월 24일 23시 59분

이번 숙제의 목표는 다양한 training parameter가 Fashion-MNIST의 학습 정확도 및 학습 시간에 끼치는 영향을 확인하고, training 과정에서의 weight와 accuracy 경향을 이해하는 것이다. 문제 1, 2, 3번을 위하여 아래와 같은 코드가 주어진다.

- prob1.py: 문제 1, 2, 3번 실험을 위한 baseline training 코드.

구현 및 실험 결과에 대한 보고서를 작성해야 한다. 구현해야 하는 내용과 보고서에 포함되어야 하는 내용은 아래와 같다.

1 문제 1: Training Parameter 변화

다양한 training parameter가 학습 정확도 및 학습 시간에 끼치는 영향을 확인하라. prob1.py에서 아래와 같이 파라미터를 변경/실행하여, 각각의 경우에 대해 training 시간, training 결과 및 test accuracy를 비교하라 (총 6개의 결과).

- 주어진 Script 그대로 실행
- Batch size를 16으로 설정 / 64로 설정
- Learning rate를 0.01로 설정 / 0.0001로 설정
- Optimizer를 SGD로 변경

문제 조건에서 명시하지 않은 내용(사용하는 DNN, epoch 수 등)은 기존 script에서 주어진 그대로 사용한다. 실행 시간은 전체 실행시간(script 전체를 실행하는 시간) 과 Training 시간(모델 학습이 실제로 이루어지는 시간) 두 가지를 모두 측정한다.

2 문제 2: Gradient 변화

Baseline 코드에서 1번째 layer와 2번째 layer 사이의 gradient의 squared sum을 출력하는 코드를 작성하고, epoch 수가 증가함에 따라 gradient의 squared sum이 어떻게 변하는지 그래프로 도시하고, 그 이유를 설명하라. 아래의 Baseline 코드에서 1번째 layer는 `nn.Linear(784, 128)`이고, 2번째 layer는 `nn.ReLU()`이다. (cf. 문제에서 말하는 1번째 layer와 2번째 layer 사이의 gradient는 강의 노트 Lecture 15의 24 page에서 $\frac{\partial E}{\partial s_j^l} := \delta_j^l$ 를 의미한다.)

```
model = nn.Sequential(nn.Linear(784, 128),
                      nn.ReLU(),
                      nn.Linear(128, 10),
                      nn.LogSoftmax(dim = 1)
                      )
```

3 문제 3: Gradient Initialization

Baseline 코드에서 gradient를 초기화하는 코드의 위치를 찾아 보고서에 기술하라. 만일 초기화를 하지 않는 경우 학습이 어떻게 되는지 직접 실행해보고, 그 결과를 첨부하여 accuracy 관점에서 비교 분석하고 그 이유를 기술하라.

4 구현 및 제출

ETL을 통해 보고서 파일(report.pdf)을 제출한다.

파일	설명
report.pdf	실험 방법 및 결과가 기술된 보고서(ETL을 통해 제출)

이번 과제는 실험 결과 보고서만 제출하도록 한다. 보고서에는 아래와 같은 내용이 포함되어야 한다.

- 실험 환경 (실습 서버 및 Jupyterhub를 사용한 경우 명시하지 않아도 됨)
- 각각의 실험을 위해 기존 script에서 어느 부분을 수정하였는지 설명
- 성능 측정을 위해 기존 script에서 어느 부분을 수정하였는지 설명
- 주어진 코드가 실제로 GPU를 사용하는지 확인한 방법 설명
- 문제 1, 2, 3번 실험 결과

제출 시 아래와 같은 사항을 유의하라.

- 이번 과제에서는 원하는 대로 자유롭게 소스 코드를 수정/모듈을 import하여 사용해도 된다. 단, 보고서에 사용한 module 및 소스코드 수정 내역을 명시하여야 한다.
- 반드시 GPU를 사용할 수 있는 환경에서 실험을 진행해야 한다. 로컬 PC에 환경이 설정되어 있으면 로컬 PC에서 실험을 진행하면 되며, GPU 환경이 구비되지 않은 경우 실습 클러스터의 JupyterHub 환경을 사용하여 실험을 진행하면 된다 (Single GPU instance 사용).
- 이번 과제는 ETL을 통해 제출하도록 한다. 실습 서버에 보고서를 제출하더라도 숙제 제출을 받지 않는다.
- 그레이스 데이를 사용할 경우 spdsta@aces.snu.ac.kr로 이를 알려야 하며, 메일 제목은 [SPDS] HW6-Graceday 학번 이름의 포맷으로 작성하라 (예: [SPDS] HW6-Graceday 2020-20000 안규수).