강화학습

Q-Network - Quiz

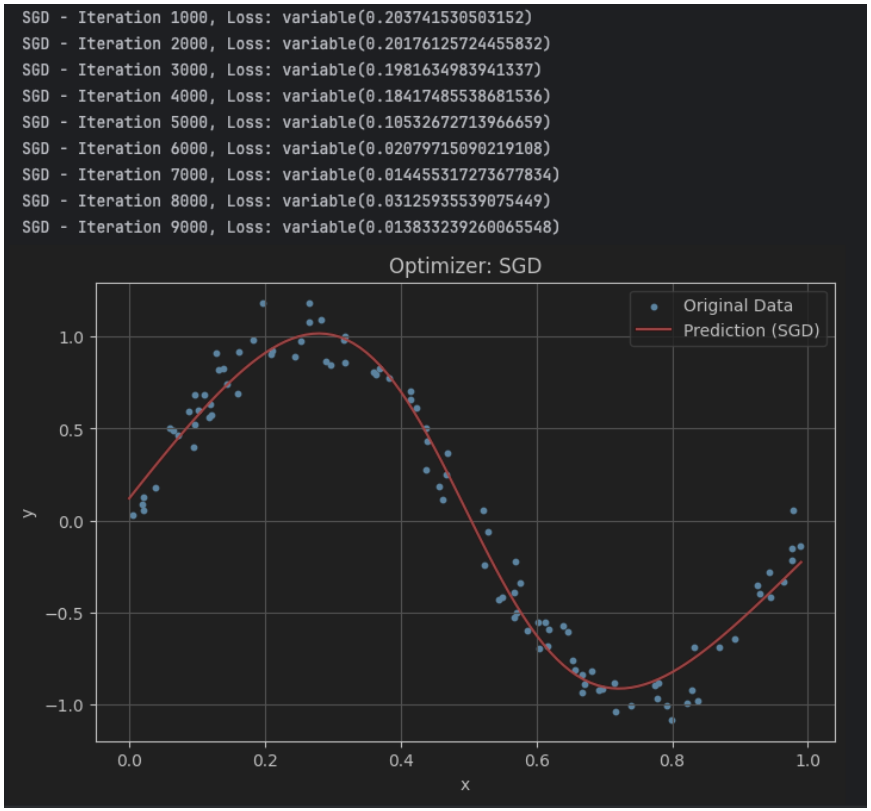
학과 : 산업인공지능학과
학번 : 2024254022
이름 : 정현일

2025.05.01.

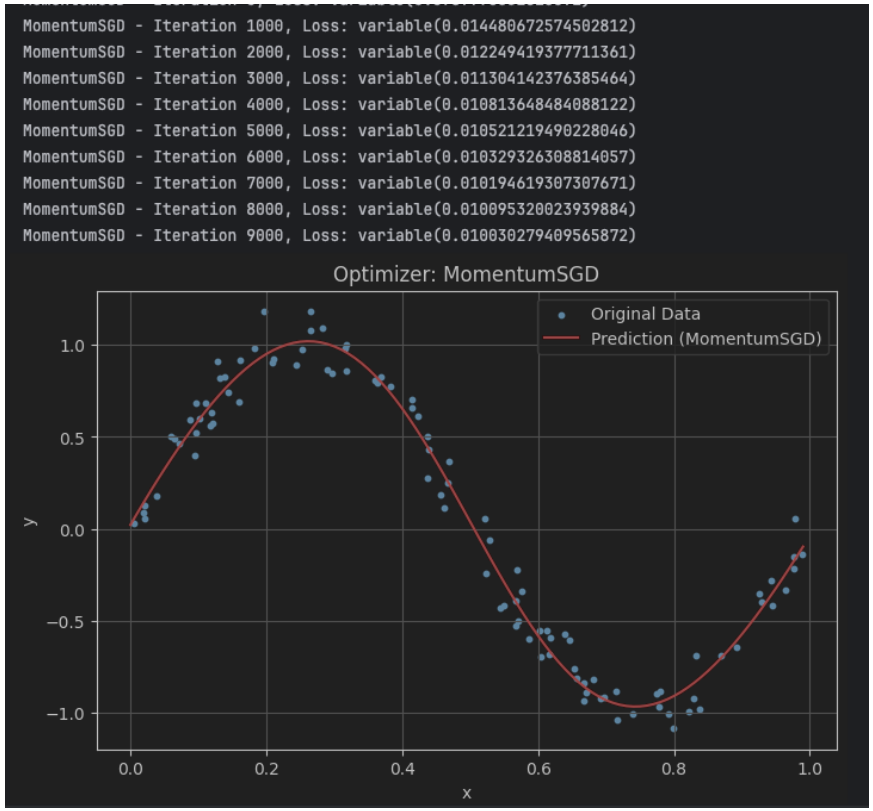
충북대학교 산업인공지능학과

(Q1) 위 실습에서 다음의 optimizer 를 실행하여 결과를 비교하라.

```
class MomentumSGD(Optimizer):
class AdaGrad(Optimizer):
class Adam(Optimizer):
```



SGD

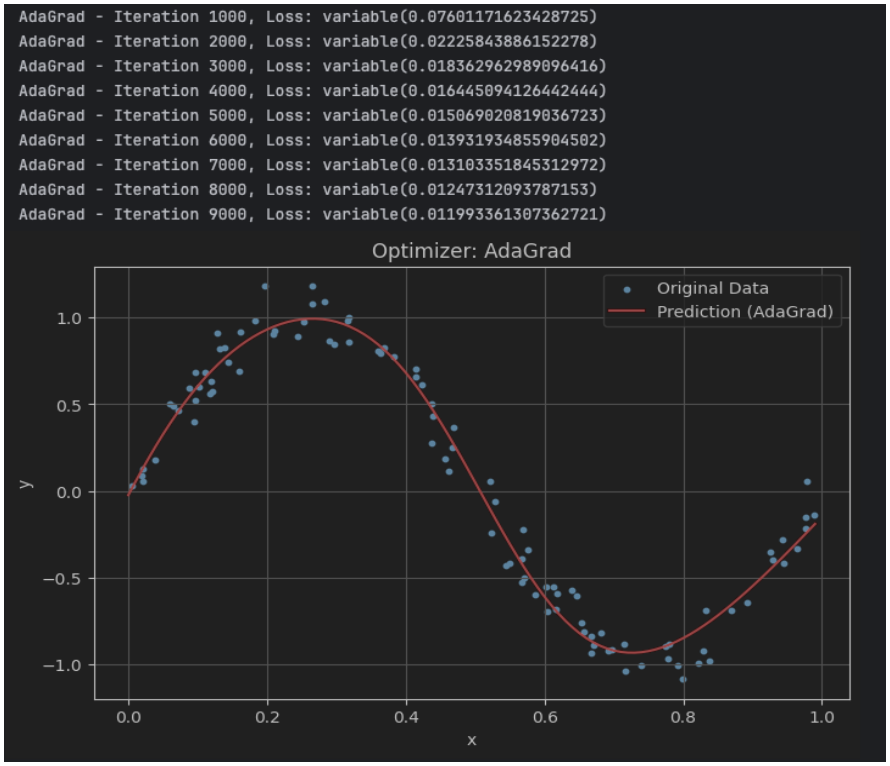


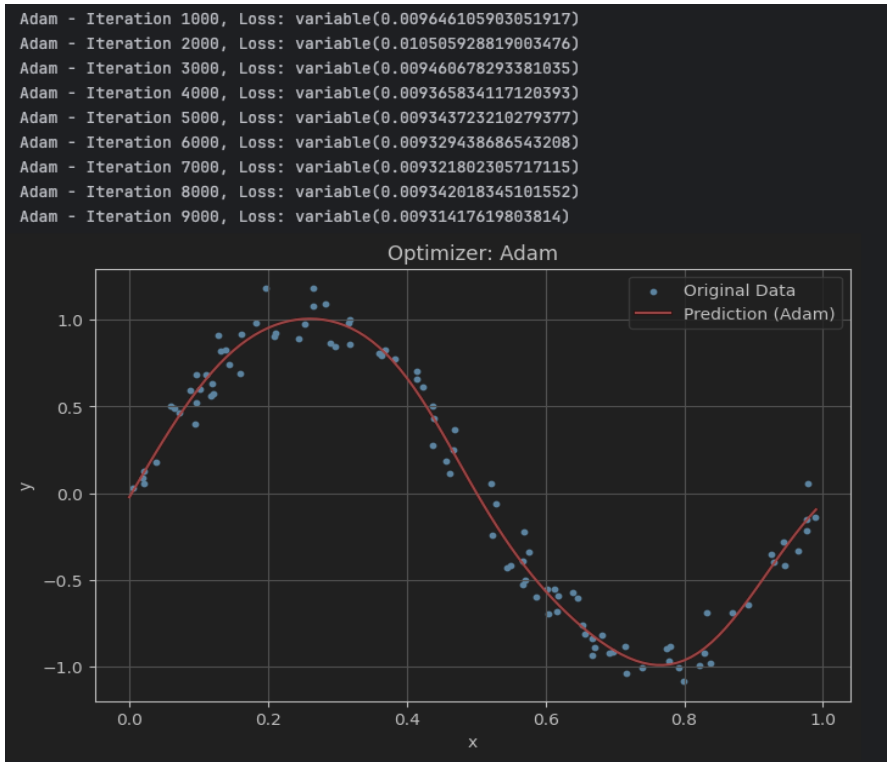
MomentumSGD

강화학습

(Q1) 위 실습에서 다음의 optimizer 를 실행하여 결과를 비교하라.

Optimizer 모드 각각 시험해본 결과 Adam 0.009314176198038l4 으로 제일 적은 loss값을 나타낸다





AdaGrad

SGD: 0.013833239260065548

AdaGrad: 0.011993361307362721

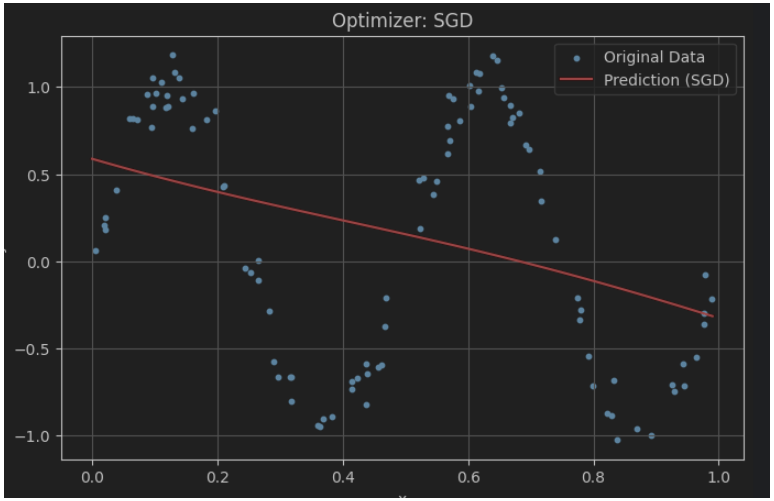
Adam

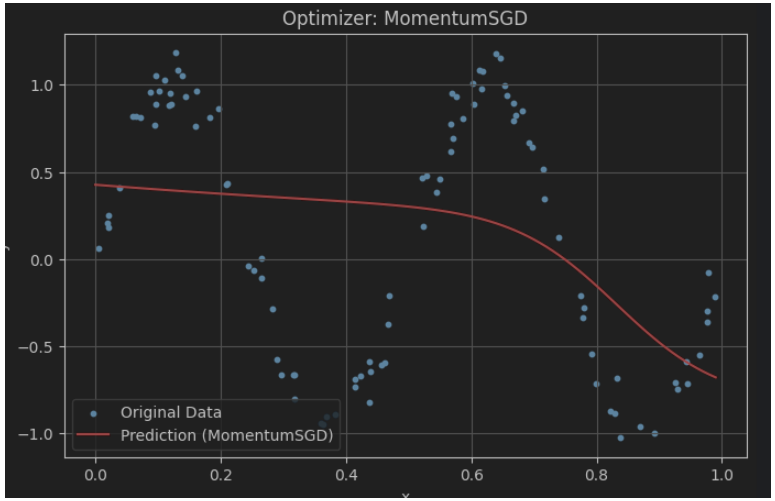
MomentumSGD: 0.010030279409565872

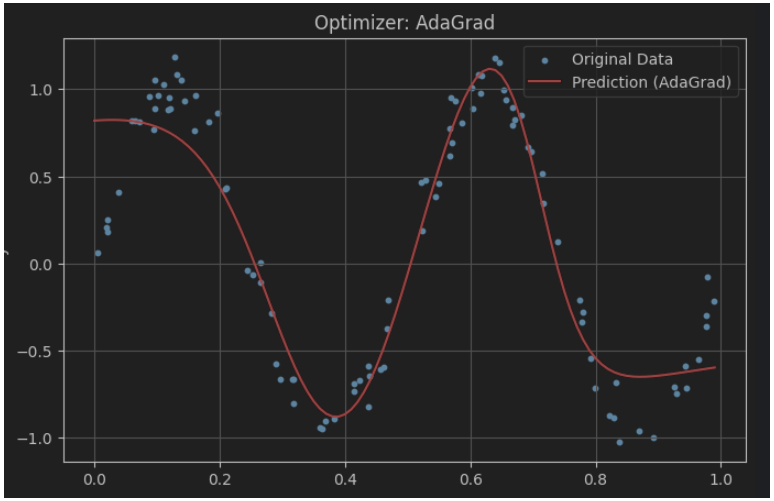
Adam: 0.009314176198038l4

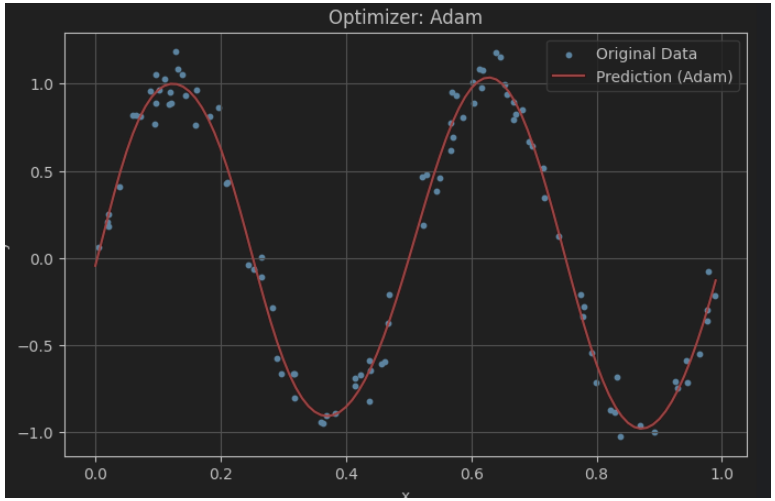
강화학습

(Q2) $y = \sin(4\pi x)$ $(0 \le x \le 1)$ 에 대한 loss 가 최소화 되도록 신경망을 최적화하고 결과를 출력하라.









강화학습

(Q3) Q-Network 를 적용하여 5x5 Grid World 에 대한 Q 테이블을 완성하고 policy 를 구하라. 단, 신경망의 최적화를 위한 파라메터를 설정하라.

[Q테이블]

```
State: (0, 0), Action: 0, Q-value: 0.5082912445068359
State: (0, 0), Action: 1, Q-value: 0.4714983105659485
State: (0, 0), Action: 2, Q-value: 0.4341954290866852
State: (0, 0), Action: 3, Q-value: 0.8370070457458496
State: (0, 1), Action: 0, Q-value: 0.6570677161216736
State: (0, 1), Action: 1, Q-value: 0.6117208003997803
State: (0, 1), Action: 2, Q-value: 0.6710969805717468
State: (0, 1), Action: 3, Q-value: 0.8939908146858215
State: (0, 2), Action: 0, Q-value: 0.8971067667007446
State: (0, 2), Action: 1, Q-value: 0.7862273454666138
State: (0, 2), Action: 2, Q-value: 0.7411519885063171
State: (0, 2), Action: 3, Q-value: 0.9715120196342468
State: (0, 3), Action: 0, Q-value: 0.5105817914009094
State: (0, 3), Action: 1, Q-value: 0.45226749777793884
State: (0, 3), Action: 2, Q-value: 0.45922234654426575
State: (0, 3), Action: 3, Q-value: 0.580070972442627
State: (1, 0), Action: 0, Q-value: 0.5291873812675476
State: (1, 0), Action: 1, Q-value: 0.5850884318351746
State: (1, 0), Action: 2, Q-value: 0.2875347137451172
State: (1, 0), Action: 3, Q-value: 0.3879384994506836
State: (1, 1), Action: 0, Q-value: 0.7368839979171753
State: (1, 1), Action: 1, Q-value: 0.7735257148742676
State: (1, 1), Action: 2, Q-value: 0.8471799492835999
State: (1, 1), Action: 3, Q-value: 0.7462244629859924
State: (1, 2), Action: 0, Q-value: 0.8802248239517212
State: (1, 2), Action: 1, Q-value: 0.7530813217163086
State: (1, 2), Action: 2, Q-value: 0.8204984068870544
State: (1, 2), Action: 3, Q-value: -0.14373673498630524
State: (1, 3), Action: 0, Q-value: 1.033362865447998
State: (1, 3), Action: 1, Q-value: 0.6747794151306152
State: (1, 3), Action: 2, Q-value: 0.7625390291213989
State: (1, 3), Action: 3, Q-value: 0.3932201862335205
State: (2, 0), Action: 0, Q-value: 0.5319892764091492
State: (2, 0), Action: 1, Q-value: 0.5564687252044678
State: (2, 0), Action: 2, Q-value: 0.591364860534668
State: (2, 0), Action: 3, Q-value: 0.6222833395004272
State: (2, 1), Action: 0, Q-value: 0.6198688745498657
State: (2, 1), Action: 1, Q-value: 0.5100265145301819
State: (2, 1), Action: 2, Q-value: 0.5905502438545227
State: (2, 1), Action: 3, Q-value: 0.6299696564674377
State: (2, 2), Action: 0, Q-value: 0.7750536799430847
State: (2, 2), Action: 1, Q-value: 0.5740088820457458
State: (2, 2), Action: 2, Q-value: 0.6239744424819946
State: (2, 2), Action: 3, Q-value: 0.6203243136405945
State: (2, 3), Action: 0, Q-value: -0.00357205793261528
State: (2, 3), Action: 1, Q-value: 0.372499018907547
State: (2, 3), Action: 2, Q-value: 0.722237765789032
State: (2, 3), Action: 3, Q-value: 0.3673689365386963
```

[Policy]

```
State: (0, 0), Best Action: RIGHT
State: (0, 1), Best Action: RIGHT
State: (0, 2), Best Action: RIGHT
State: (0, 3), Best Action: RIGHT
State: (1, 0), Best Action: DOWN
State: (1, 1), Best Action: LEFT
State: (1, 2), Best Action: UP
State: (1, 3), Best Action: UP
State: (2, 0), Best Action: RIGHT
State: (2, 1), Best Action: RIGHT
State: (2, 2), Best Action: UP
State: (2, 3), Best Action: LEFT
```

강화학습

(Q3) Q-Network 를 적용하여 5x5 Grid World 에 대한 Q 테이블을 완성하고 policy 를 구하라. 단, 신경망의 최적화를 위한 파라메터를 설정하라.

[실행결과]

```
Episode 900/1000, Loss: 0.0035, Epsilon: 0.1000
```