

hidden_size 와 learning_rate가 학습에 미치는 영향

2024254012 배인호

교재 p.143에 있는 프로그램을 수정하여 다음의 경우에 대한 학습 결과 그래프를 출력하시오

- (1) hidden_size=10으로 하였을때
- (2) hidden_size=100으로 하였을때
- (3) learning_rate= 0.01일때
- (4) learning_rate=0.5 일때의

위의 시뮬레이션 결과를 이용하여, hidden_size 와 learning_rate가 학습에 미치는 영향을 기술하시오

1. hidden_size의 영향 : 은닉층의 크기 뉴런의 수

가. Hidden layer의 뉴런 수를 작게 할 때(hidden_size=10), 모델이 충분한 표현력을 가질 수 없어서 학습이 덜 될 수 있다. 표현력이 부족하면 데이터의 복잡한 패턴을 학습하지 못하고 underfitting 발생 가능성 있음. hidden_size를 작게 설정하면 성능이 좋지 않을 가능성이 있다.

나. Hidden layer의 뉴런 수를 크게 할 때(hidden_size=100), 모델이 보다 복잡한 패턴을 학습할 수 있다. 그러나 hidden_size가 커질수록 모델의 복잡성이 증가하기 때문에, 과적합(overfitting) 문제가 발생할 수 있음. hidden_size를 너무 크게 설정하면 훈련 데이터에 너무 맞춰져 새로운 데이터에 대한 일반화 성능이 떨어질 가능성이 있다.

2. learning_rate의 영향 : 학습률 가중치 업데이트 적용 비율

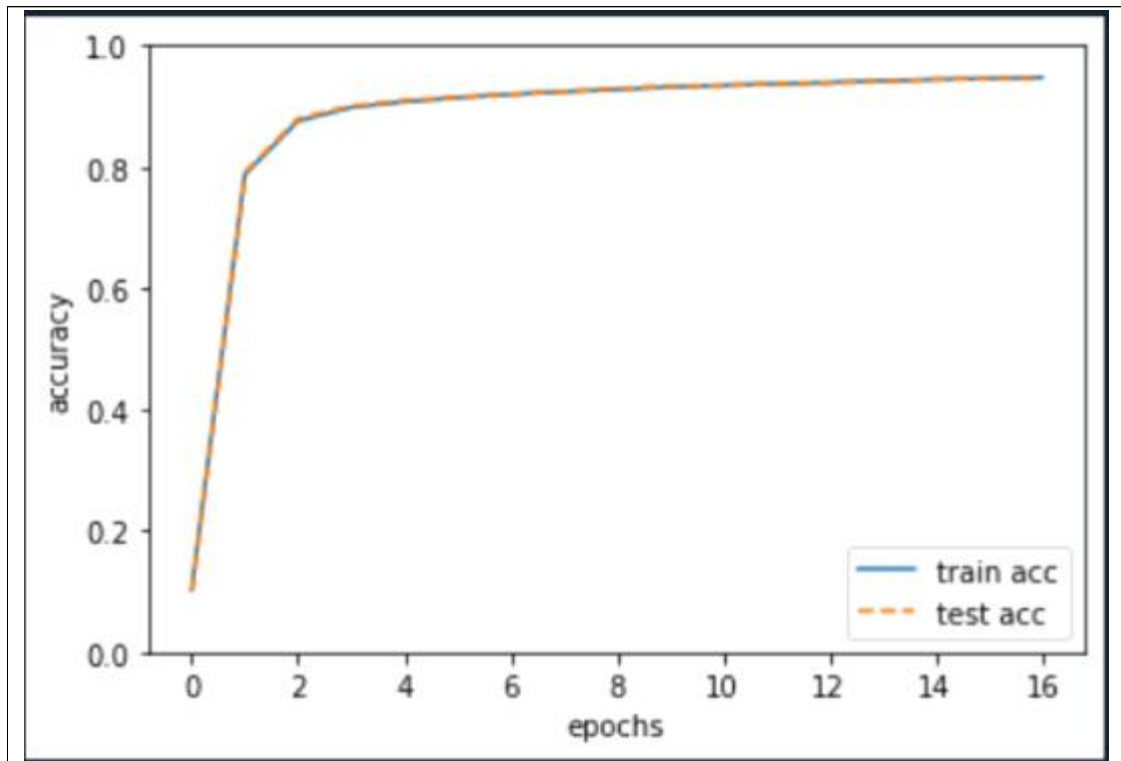
가. 학습률이 작을 때(learning_rate=0.01), 매개변수 갱신이 조금씩 이루어지므로 학습 속도가 느릴 수 있다. local minimum에 빠져서 학습이 잘 이루어지지 않을 가능성 있음. 너무 큰 학습률보다는 안정적으로 학습이 진행될 가능성이 높습니다.

나. 학습률이 클 때(learning_rate=0.5), 매개변수 갱신이 빠르게 이루어지므로 학습 속도가 빠를 수 있다. 그러나 학습률이 크면 최적의 해(solution)를 지나칠 수 있어서 학습이 불안정해질 수 있음. 발산(divergence)하여 학습이 제대로 이루어지지 않을 수도 있다.

결론 : 적절한 hidden_size와 learning_rate를 선택.

hidden_size와 learning_rate를 조절하여 모델의 학습이 안정적이고 최적의 성능을 발휘할 수 있도록 하는 것이 중요함.

3. hidden_size=50, learning_rate = 0.1 일 때 기본 그래프 결과



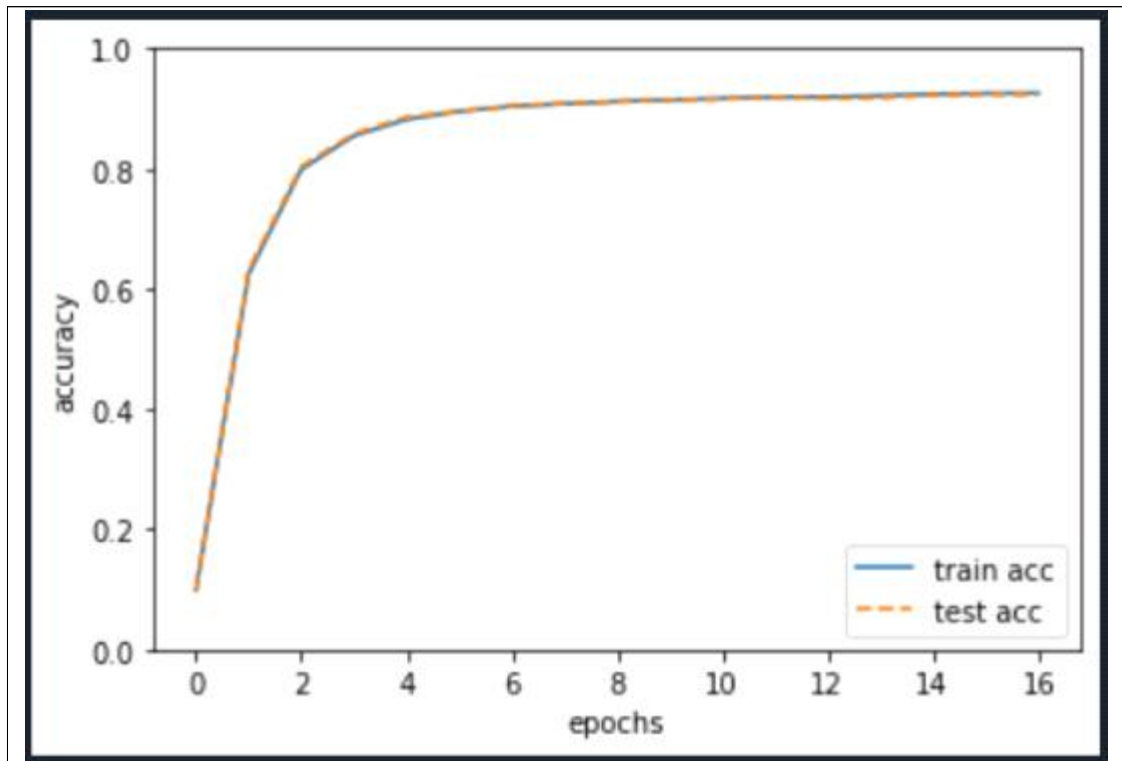
가. Hidden size (은닉층 크기) - 뉴런의 수

- 1) 은닉층이 크면 모델이 더 많은 정보를 학습 가능
- 2) 과적합 위험이 존재

나. Learning Rate (학습률) - 가중치 업데이트 적용 비율

- 1) 학습률이 높으면 빠르게 정확도 수렴
- 2) 학습 과정이 불안하다는 단점 존재

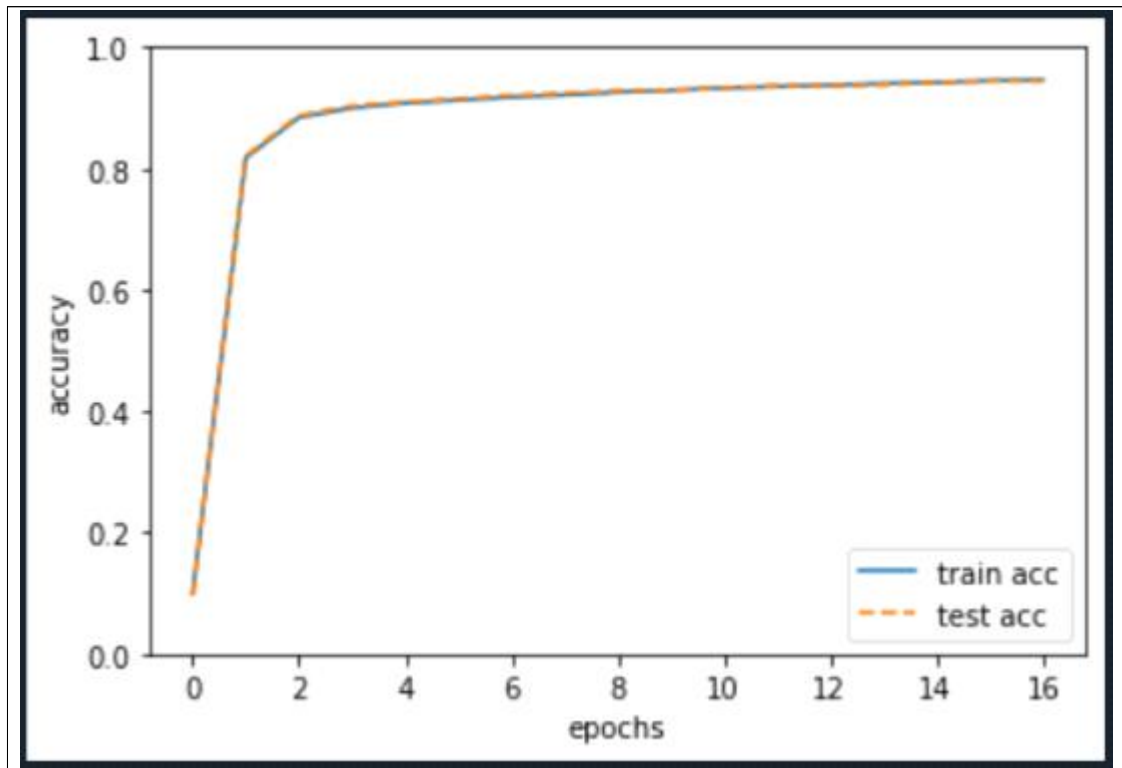
4. hidden_size=10, learning_rate = 0.1 일 때 그래프 결과



가. 은닉층 크기를 10으로 적용 학습률은 그대로

- 1) 은닉층의 크기를 줄였기 때문에 데이터의 복잡성을 충분히 학습을 못해 훈련 정확도가 낮을 가능성이 있음
- 2) 과소적합이 될 가능성도 있음.
- 3) 학습이 잘 이루어지지 않을 가능성도 존재 (local minimum)

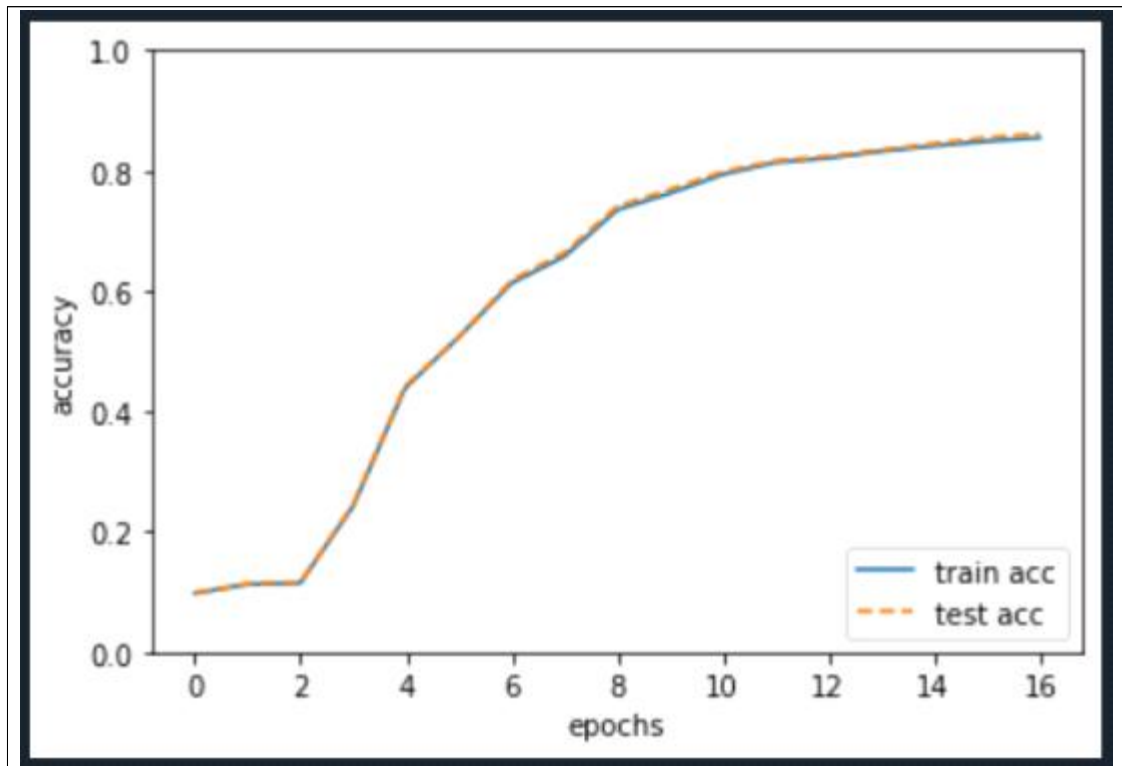
5. hidden_size=100, learning_rate = 0.1 일 때 그래프 결과



가. 은닉층 크기를 100으로 학습률은 그대로

- 1) 모델이 더 많은 정보를 학습을 하기 때문에 정확도가 빠르게 높아지는 현상을 보임
- 2) 정확도는 높게 형성이 되지만 과적합의 우려가 있을 수 있음
- 3) 은닉층의 크기를 너무 크게 하게 되면 일반화 성능이 떨어질 수도 있음.

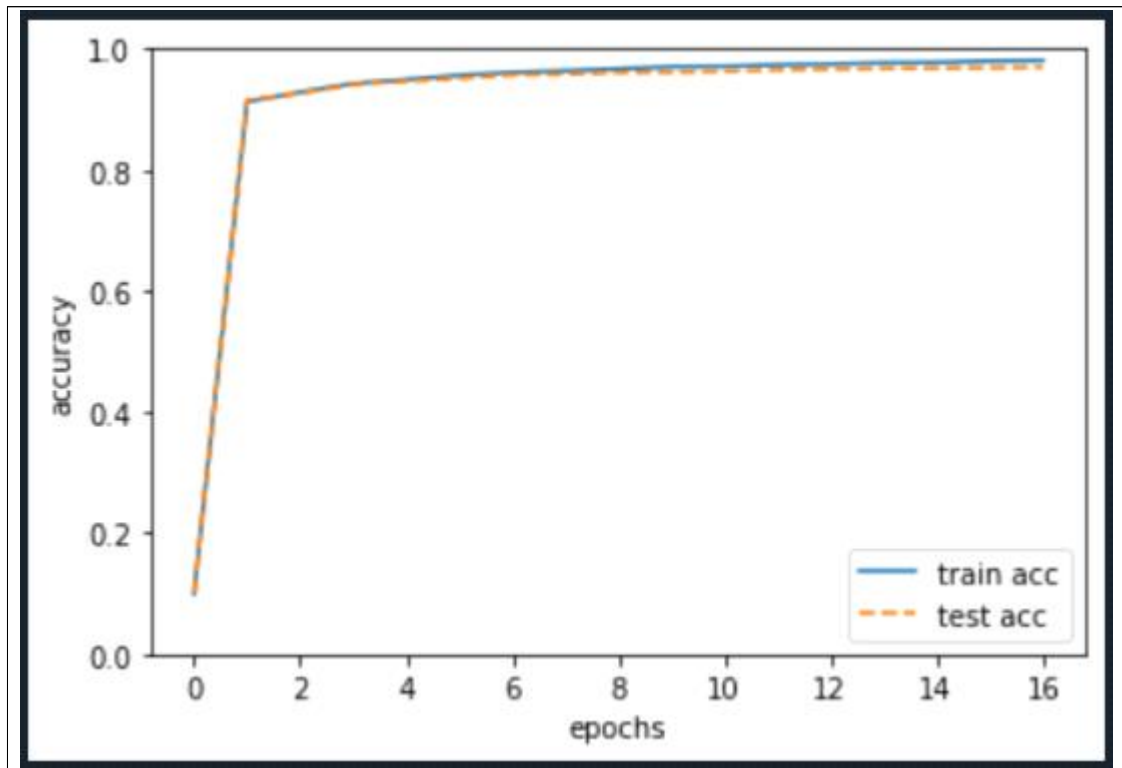
6. hidden_size=50, learning_rate = 0.01 일 때 그래프 결과



가. 은닉층의 크기는 그대로 두고 학습률을 낮게 설정할 경우

- 1) 변수의 업데이트가 작기 때문에 정확도가 천천히 증가하는 현상
- 2) 학습과정은 안정적이고 더 나은 결과값을 얻을 수 있음
- 3) 학습속도가 느리다는 단점을 가지고 있음.

7. hidden_size=50, learning_rate = 0.5 일 때 그래프 결과



가. 은닉층의 크기를 그대로 두고 학습률을 높게 설정할 경우

- 1) 학습을 빠르게 진행하여 정확도가 크게 변동
- 2) 학습이 불안정하여 최적값을 지나칠 가능성을 가지고 있음
- 3) 손실이 빠르게 감소하지만 불안정하게 변동하거나 증가할 수 있음.