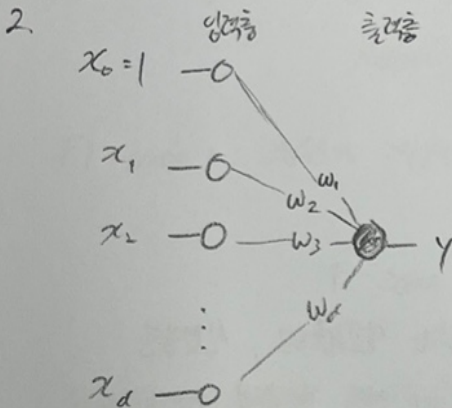


공간고사 문제풀이

2023054023 강지호

1. 모델이 더 큰 클래스의 패턴을 과하게 학습하여 소수 클래스에 대해서는 일반화 성능이 떨어지는 과적합이 발생할 수 있다. 이를 해결하기 위해서는 데이터 양을 늘리거나 모델의 복잡도를 줄이는 등의 방법을 사용할 수 있다.



피생크의 구조는 입력층과 출력층이라고 부르는
2개의 층이 있고, 입력층은 여러 연산을 하지 않아
층의 개수를 세지 않는다.

3. 전이 학습은 이미 한 작업에서 학습된 모델을 기반으로 하여, 다른 작업에 적용하고 이를 통해 학습 과정을 가속화하거나 학습 효율을 향상 시키는 기법이다. 기본적으로, 이미 대량의 데이터로 훈련된 모델의 지식을 이용하여, 새로운 비슷한 작업에 대한 학습을 보다 빠르고 효율적으로 수행하도록 한다.
4. CNN 학습을 처음에서는 특징을 식별하여 이미지 전체에 걸쳐 공유된다. 이를 통해 모델의 파라미터 수를 크게 줄이고, 학습이 더 효율적으로 이루어진다. 또한 CNN은 이미지의 코어 패턴부터 시작하여 점점 더 복잡하고 추상적인 고수 특징으로 점차 변환한다. 이를 통해 객체의 공간적 구조를 효과적으로 학습한다. 이미지 내의 객체가 위치를 바꿔도, CNN은 해당 객체를 인식할 수 있다. 이것이 이미지 전체를 스캔하여 가능하기 때문이다.

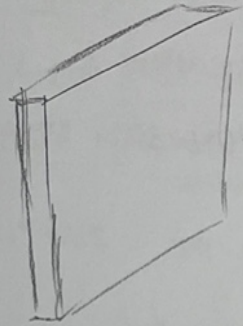
학습용 데이터 : 이미지에서 특징을 추출하는 극초기 계층.

활성화 함수 : 대부분의 CNN에서는 ReLU 함수를 함수가 사용된다.

풀링 계층 : 전체 이미지의 크기를 줄이거나 포착하여 지역성을 감소시키고, 앞의 층에 대한 불변성을 제공한다.

완전 연결 계층 : 마지막에 위치하여, 앞서 추출된 특징들을 바탕으로 이미지를 분류하거나 데이터셋을 예측한다.

5.



$$I_h \times I_w \times I_c$$

*



$$K_h \times K_w$$

*



$$O_h \times O_w \times O_c$$

$$\Rightarrow \text{총 개수 } s = K_h \times K_w \times I_c \times O_c$$

6. 2차 역전파는 신경망에서 가중치를 효율적으로 조정해 주며 사용되는 알고리즘으로, 신경망을 통해 입력의 전역에 출력층까지 도달한 후, 실제 출력과 기대 출력 사이의 오차를 계산한다. 이 오차를 역으로 계산하기 위해 2차로 입력층으로 되돌려 보내면서 각 층의 가중치를 조정하는 과정이다.

가중치 소실 문제는 실용 신경망에서 발생하는 중요한 현상이다. 이 문제를 해결하기 위해 역전파 과정에서 계산된 가중치가 점점 작아져 0에 가까워지는 문제를 말한다. 이를 완화하기 위해 4가지 방법이나 리미티드 레이어를 사용한다.

이를 피하기 위해 사용하는 ReLU 활성화 함수 사용하거나 가중치 초기화 기법을 개선하는 방법은 사용할 수 있다. ReLU는 많은 입력에 대해 0을 출력하고, 양의 입력에 대해서 입력을 그대로 출력한다.

He 초기화, Xavier 초기화 등이 이 문제의 원인을 줄여도 가중치 값을 조정해 주므로 해결할 수 있다.

7.

		Predict	
		positive	Negative
Actual	positive	TP	FN
	Negative	FP	TN

• 정밀도 : 모델이 Positive로 예측한 항목 중 실제로 Positive인 항목의 비율

$$\text{Precision} = \frac{TP}{TP + FP}$$

• 재현율 : 실제 positive인 항목 중 모델이 Positive로 정확히 예측한 항목의 비율

$$\text{Recall} = \frac{TP}{TP + FN}$$

• F1-score : 정밀도와 재현율의 조화평균.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

8. • 확률적 경량 방법 : 각 반복에서 각 레이어의 학습매개변수를 사용하여 가중치를 제한한다. 레이어 방법은 매우 빠른 업데이트가 가능하고 더 나은 데이터셋에 적합하다. 하지만, 높은 변동성으로 인해 학습 과정이 불안정할 수 있고, 최적점까지 수렴이 느릴 수 있다.

• 오엔타 : 과거 가중치의 가중치 평균을 현재 업데이트에 추가하여, 더 많은 양의 관찰을 제공한다. 이는 SGD의 변동성을 줄이고 학습 과정을 가속화한다. 더 나은 데이터셋에 적합하다. 레이어는 최적점을 넘어서는 경우가 발생할 수 있다.

9. • L1 규제 : 모델의 가중치에 대한 손실함수에 가중치의 절대값에 비례하는 항을 추가한다. 이는 일차 가중치를 정확히 0으로 만드는, 레이어 특성을 모델에서 제거시키는 효과가 있다.

$$L1 = \gamma \sum |w|$$

• L2 규제 : 손실함수에 가중치의 제곱에 비례하는 항을 추가하여, 가중치의 크기를 제한한다. 이는 모든 가중치를 작게 유지하여 모델의 복잡도를 줄인다.

$$L2 = \gamma \sum w^2$$