

1. 불균형 데이터셋의 문제점과 함께 그것을 해결하는 방법에 대해 설명하시오

문제점 - 불균형 데이터셋은 특정 클래스의 데이터가 다른 데이터 클래스에 비해 현저히 많거나 적은 경우를 말하는데, 이로 인해 모델이 다수 클래스에 편향되기 학습될 수 있으며, 결과적으로 소수 클래스의 정확도의 정확도가 낮아지는 문제가 생긴다.

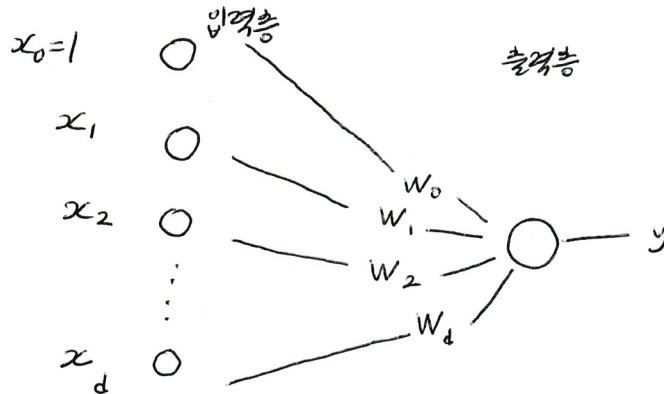
해결방법 - C버 샘플링 : 소수 클래스의 데이터를 증강시켜 불균형을 해소한다

언더 샘플링 : 다수 클래스의 데이터를 줄여 불균형을 해소한다

SMOTE : 소수 클래스의 데이터를 인공적으로 생성하여 C버 샘플링한다.

가중치 부여 : 소수 클래스 데이터의 손실함수에 더 큰 가중치를 부여하여 불균형을 해소한다.

2. 퍼셉트론의 구조를 그림으로 설명하시오



3. 전이 학습에 대해 설명하시오

전이 학습은 사전에 훈련된 모델을 새로운 문제에 적용하는 학습 방법이다. 이는 큰 데이터셋에서 학습된 모델의 지식을 작은 데이터셋 문제에 전이하여 4통항으로써 학습 시간을 단축하고, 작은 데이터셋에서도 효율적인 학습이 가능하게 한다.

4. CNN 아키텍처의 특징과 주요 구성요소별 구조를 서술하시오.

특징 - CNN은 이미지 인식과 분류에 탁월한 성능을 보이는 신경망 구조이다. CNN은 지역적인 패턴을 인식하는 능력이 뛰어나며, 파라미터의 수가 상대적으로 적어 과적합을 방지할 수 있다.

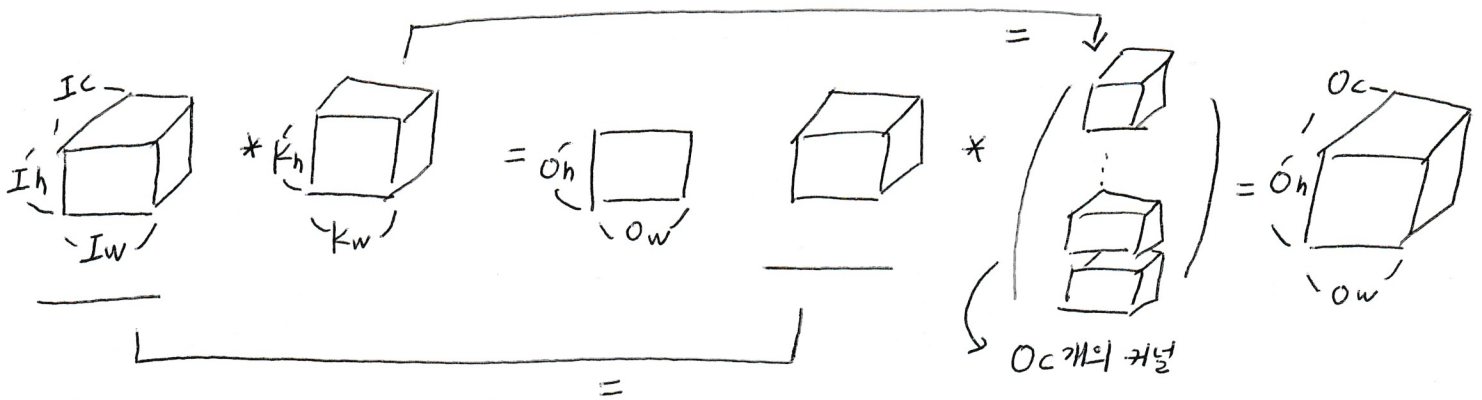
주요 구성요소 - 합성곱 층 (Convolutional Layer) : 이미지에서 특징을 추출하는 층이다

풀링 층 (Pooling Layer) : 특징의 크기를 줄이거나 강하게 줄여 계산량을 줄이는 역할을 한다

완전 연결 층 (Fully Connected Layer) : 합성곱 층과 풀링 층을 통해 추출된 특징들을 기반으로 최종적으로 분류를 수행한다

구조 - 일반적으로 여러 개의 합성곱 층과 풀링 층이 번갈아 나타나며, 마지막에는 하나 또는 여러 개의 완전 연결 층이 위치한다.

5. Convolution 연산의 가중치 parameter의 총 수를 계산하는 과정을 그림으로 설명하시오



6. C와 역전파의 개념과 가중치 소멸 문제를 설명하고, 이에 대한 해결방법을 설명하시오

C와 역전파 - 신경망에서 출력층에서 발생한 C를 입력층 방향으로 거슬러 올라가며 각 층의 가중치를 조절하는 학습 방법이다. 이 과정에서 체인룰을 사용하여 C와 각 층의 가중치에 대해 미분하고, 이를 통해 가중치를 업데이트 한다.

가중치 소멸 문제 - 심층 신경망에서 일어나는 현상으로, 입력층에 가까울수록 그래디언트가 점점 작아져 가중치 업데이트가 거의 이루어지지 않게 되는 문제이다. 활성화 함수나 초기 가중치 설정 등에 의해 발생할 수 있다.

해결방법 - ReLU 활성화 함수 사용: 가중치 소멸 문제를 줄일 수 있는 활성화 함수이다.  
가중치 초기화 방법 개선: He 초기화나 Xavier 초기화나 같은 선형망의 각 층에 적합한 방법을 가중치를 초기화 한다.  
배치 정규화: 각 층의 입력을 정규화하여 가중치 소멸 문제를 줄인다.  
잔여 연결: 입력을 출력에 더함으로써 그래디언트가 깊은 층까지 전달될 수 있도록 돕는다.

7. Confusion matrix를 그림으로 그리고, 정밀도, 재현율, F1-Score에 대해 수식과 함께 설명하시오

	positive	Negative	
positive	True (TP) True positive	False (FN) Negative Type II Error	sensitivity TP (TP+FN)
Negative	False (FP) positive Type I Error	True (TN) Negative	specificity TN (TN+FP)
	Precision TP (TP+FP)	Negative predict value rate TN (TN+FN)	Accuracy TP+TN (TP+TN+FP+FN)

정밀도는 모델이 양성으로 (positive) 예측한 항목 중 실제로 양성인 항목의 비율을 의미한다. True positives를 True positives와 False positives의 합으로 나눈 값이다.  $\frac{TP}{(TP+FP)}$   
재현율은 실제 양성인 항목 중 모델이 양성으로 올바르게 예측한 항목의 비율이다. True positives를 True positives와 False negatives의 합으로 나눈 값이다.  $\frac{TP}{(TP+FN)}$

F1-Score은 정밀도와 재현율의 조화 평균으로,  $F1\text{-score} = 2 \times \frac{\text{precision} \times \text{Recall}}{\text{precision} + \text{Recall}}$  로 계산된다.

8. gradient 기반 최적화 알고리즘의 종류와 2 종류에 대해 설명하시오

Gradient 기반 최적화 알고리즘은 신경망의 가중치를 조정하여 손실 함수를 최소화하기 위해 사용되는 방법이다.

SGD - 가장 기본적인 알고리즘의 최적화 형태로, 각 반복에서 임의의 데이터 샘플을 사용하여 그래디언트를 계산한다.

Momentum - SGD의 변형으로 출력을 높이기 위해 과거 그래디언트 커스 가중 평균을 사용하여 가중치 업데이트에 관성을 부여한다.

Adagrad - 각 매개변수에 대해 개별적인 학습을 적용한다.

RMSprop - Adagrad의 단점을 해결하기 위해 과거 모든 그래디언트를 고려하는 대신, 작은 그래디언트에 더 많은 가중치를 주어 계산한다.

Adam - Momentum과 RMSprop의 개념을 결합한 알고리즘으로, 모멘텀을 사용하여 경사 하강법의 방향을 유지하면서 각 매개변수에 대한 적응형 학습률을 적용한다.

9. 학습 시 overfitting을 방지하기 위한 규제 기법에 대해 설명하시오

overfitting을 방지하기 위한 규제 기법은 모델이 학습 데이터에만 지나치게 최적화되는 것을 방지하고, 일반화 능력을 향상시키기 위해 사용된다.

L1 규제 - 모델 가중치의 절대 값에 비례하는 비용을 손실 함수에 추가하여, 일부 가중치를 0으로 만들어 모델을 단순화한다.

L2 규제 - 가중치의 제곱에 비례하는 비용을 손실 함수에 추가하여 가중치 값을 감소시키고, 모델의 복잡도를 줄인다.

Dropout - 학습 중에 무작위로 뉴런을 비활성화하여 모델이 특정한 뉴런이나 패턴에 과도하게 의존하는 것을 방지한다.

Early stopping - 학습 과정에서 검증 데이터셋의 성능이 개선되지 않을 때 학습을 조기 중단하여 과적합을 방지한다.