**NN\_Adv\_Assignment\_1**

학습률은 딥러닝 모델 학습 과정에서 매우 중요한 하이퍼파라미터 중 하나다. 학습률은 각 반복마다 가중치를 얼마나 조정할지를 결정하고, 이 값이 너무 크면 발산할 수 있고, 너무 작으면 학습이 느려진다. 이를 극복하기 위해 학습률 스케줄러를 사용한다. 학습률 스케줄러는 학습 과정 중 학습률을 조절하여 최적의 학습률을 찾는데 도움을 준다. 다양한 학습률 스케줄러가 존재하지만 <1. Step Decay>방식과 <2. Exponential Decay>, <3. One-Cycle Learning Rate Policy>방식을 소개해보고자 한다. 먼저 <Step Decay>방식은 학습률을 일정한 간격마다 감소시키는 방법이다. 예를 들어 매 10,000번의 반복마다 학습률을 1/2로 줄일 수 있따. 초기에는 빠르게 학습하다가 점진적으로 작은 변화를 보여주며 안정적으로 수렴하게 된다. 그러나 이 방식은 최적 학습률을 찾는 시간이 상대적으로 더 오래걸린다는 단점이 있다. 두 번째로 <Exponential Decay>방식은 지수 함수로 줄이는 방법이다. 일반적으로 매 반복마다 학습률을 현재 학습률에 일정한 감쇠 상수를 곱하여 갱신한다. 이 방식은 초기에 빠르게 시작하고 나중에는 느리게 수렴한다. 지수 감소는 빠른 수렴을 원하면 유용하게 사용 가능하다. 마지막으로 <One-Cycle Learning Rate Policy> 방식은 총합 학습 에포크의 중간 지점에서 학습률을 증가시키고 다시 감소시키는 방식이다. 이는 초기에 빠르게 학습하고 나중에는 학습률을 낮추어 정확도를 높이는데 도움이 된다.

딥러닝 모델을 개발할 때 훈련 데이터의 오차를 최소화하면서 일반화 오차를 관리하는 것은 중요한 과제다. 두 오차 간의 간극을 줄이기 위한 4가지 방법을 소개하겠다. <1. 교차 검증(Cross-Validation)>이다. 교차 검증은 데이터를 여러 부분으로 나눈 후, 각 부분을 번갈아가며 검증 세트로 사용하여 모델을 평가하는 방법이다. 이를 통해 모델의 일반화 성능을 더 정확하게 평가할 수 있다. <2. 더 나은 최적화 알고리즘 선택>은 SGD, Adam, RMSProp 등 다양한 최적화 알고리즘을 사용하여 훈련 과정을 개선할 수 있다. 각 알고리즘에 따라 모델의 수렴 속도와 최적화 품질에 영향을 줄 수 있음은 유념해야 한다. <3. 가중치 초기화 방법 변경>을 통해 모델의 초기 상태를 조절할 수 있다. 잘못된 초기화는 모델의 수렴을 방해할 수 있으며, 이를 개선하여 빠른 수렴과 더 낮은 일반화 오차를 달성할 수 있다. 마지막으로 <4. 적절한 네트워크 아키텍처 선택>이다. 모델의 아키텍처를 선택할 때, 문제의 복잡성과 데이터의 특성에 맞는 구조를 고려하는 것이 중요하다. 너무 복잡한 구조는 훈련 오차를 줄일 수 있지만 과적합도 유발할 수 있기 때문이다. 따라서 적절한 아키텍처를 선택하여 일반화 오차를 개선할 수 있다. 대표적인 아키텍쳐는 신경망, 합성곱 신경망, 순환 신경망, 장단기 메모리, 변환자, 생성적 적대 신경망 등이 존재한다.