**인공지능개론 정리노트#2**

201904022 김상옥

202284026 안정빈

**질문1. 미니 배치의 크기가 속도와 정확도에 어떤 영향을 미치며, 적절한 미니 배치의 수는 무엇일까?**

**김상옥 :** 미니 배치 수가 크면 정확도가 더 높아지지 않을까?

**안정빈 :** 미니 배치 수가 크면 학습 속도가 빨라지니까 많은 데이터로 계산을 하면서 정확도를 높일 수 있을 것 같다는 생각이 들기는 하지만 과적합 위험도가 있을 것 같아.

**김상옥 :** 미니 배치가 범용성을 위한 학습이니까 너무 많은 크기로 학습하면 범용성이 떨어지는 게 맞는 거 같다. 그런데 학습 속도는 오히려 오래 걸리지 않을까? n번 반복하여 학습한다고 할 때 한번에 많은 양을 n번 처리하는 것 보다 적은 양을 n번 처리하는 게 더 빠르다고 생각해

**안정빈 :** 학습 속도는 미니배치 크기가 작을 경우가 더 빠른 게 맞는 것 같아! 그러면 미니배치의 크기가 클 때는 업데이트 당 속도가 느리고 많은 데이터가 있으면 메모리 부족의 위험도는 없을까?

**김상옥** : 더 복잡하고 어려운 문제를 위해서는 신경망의 깊이나 가중치의 수가 엄청 많을 것이니 분명 메모리나 하드웨어에 부담이 많을 것 같아. 그래서 엔비디아같은 회사에서 AI를 위한 하드웨어를 개발하는 거라고 생각해. 아직은 공부를 위한 간단한 데이터로 활용했지만 이마저도 생각보다 오래걸린다고 느껴졌던 걸 생각하면 실제로 사용될 AI학습은 하드웨어에 영향을 받고 시간도 많이 걸릴 것 같아. 그렇기 때문에 속도도 느리지 않으면서 정확도(범용성)를 챙길 수 있는 미니 배치의 수가 중요하다고 생각해. 아까 미니 배치의 수가 크면 범용성이 떨어질 것이라 했는데 어떤 영향이 있을까?

**안정빈 :** 우선 학습 속도가 느려지고 미니 배치 수가 크면 모델이 학습 데이터의 특정 패턴에 지나치게 의존하게 되어 일반화 능력이 저하되고 다양한 패턴을 덜 학습하게 되는 것 같아. 그래서 범용성을 고려할 때 적절한 미니 배치 크기를 선택해야 되는 것 같아.

**김상옥 :** 직접 해보니 일정 수 이상으로는 큰 의미가 있을 정도로 정확도가 개선되지는 않고, 오히려 학습 속도에 영향을 주었는데 확실히 적절한 배치 크기를 설정하는게 중요한 거 같다. 또 학습할 신경망에 따라 반복수, 미니 배치 수 등 다양한 요소를 고려하여 최선의 학습을 구하는게 AI에 중요한 점인 거 같아. **(하단에 사진 첨부 하였습니다.)**

**안정빈 :** 우리 둘다 미니 배치 크기가 적절해야 한다고 생각하는 것 같아. 그러면 적절한 미니배치의 수가 무엇인지 생각해보자

**결론 :** 적절한 미니배치의 수는 학습할 신경망에 다양한 요인을 고려하여 선택하는 게 중요하다. 그렇기에 확실한 배치 수는 없고, 상황에 따라 선택을 해야합니다.

텍스트, 스크린샷, 소프트웨어, 번호이(가) 표시된 사진

자동 생성된 설명텍스트, 스크린샷, 디스플레이, 소프트웨어이(가) 표시된 사진

자동 생성된 설명**(미니배치를 6000번 했을 경우) (미니배치를 10번 했을 경우)**

**(미니배치를 100번 했을 경우)**

텍스트, 스크린샷이(가) 표시된 사진

자동 생성된 설명

**질문2.**

**질문2. 신경망의 깊이 vs 너비, 깊은 네트워크와 넓은 네트워크의 차이점은 무엇이며 어떤 문제에 어떤 구조가 적합할까?**

**안정빈 :** 신경망의 깊이는 층의 수, 너비는 한 층의 뉴런의 수를 의미해. 깊은 네트워크는 우선 비선형이고, 구조는 단순해 보일 것 같지만 계산량은 많은 것 같아. 넓은 네트워크는 가중치가 늘어나니까 관계가 복잡할 것 같아.

**김상옥 :** 너비가 크다는 건 고려할 요인이 많다는 것 같아. 너비가 뉴런의 수이니 어떤 문제에 대해서 더 고려할 요인들(가중치)이 많으며, 그 말은 행렬의 크기가 크다는 의미이니 어떤 문제의 크기를 나타내는 게 아닐까. 그리고 깊이가 깊다는 건 그만큼 복잡하고 어려운 문제를 나타내는 것 같아. 각 활성화 함수를 통해 다음 층으로 넘어가는데 그 층이 많다는 건 그만큼 복잡하고 어려운 문제를 해결하기 위해 계산한다고 생각해.

**안정빈 :** 깊은 네트워크는 층이 많아지니까 복잡한 패턴을 학습하는데 사용 될 것 같아. 적은 매개변수로 복잡한 함수를 표현 할 것 같은데 예를들어 이미지, 오디오등 복잡한 처리 에 사용될 것 같고, 넓은 네트워크는 각 층에서 많은 뉴런을 가지고 많은 매개변수가 있어서 더 복잡한 함수를 표현 하는데 자연어처리, SNS의 대규모 데이터 베이스 같은 대용량 처리를 다룰때에도 유용 할 것같아.

**김상옥 :** 그 말대로 문제에 따라 신경망의 너비와 깊이를 설정해야 좋은 신경망이겠다. 처리할 양이 많은 문제라면 너비를 늘리고, 복잡한 문제라면 깊이를 늘리는 식으로 신경망을 구성한다면 효율적인 신경망을 만들 수 있을 거 같아.

**안정빈 :** 나도 그렇게 생각해. 따라서 깊이와 너비는 모델의 복잡성을 결정하는 중요한 요소 같아, 적절한 깊이와 너비를 선택해서 모델이 원하는 작업을 수행할수록 있도록 하는 것이 중요하다고 생각해.

**결론 :** 깊이는 복잡한 것을 해결할 때 사용되고, 너비는 해결한 문제의 양이 많을 때 적합합니다. 문제에 따라서 깊이와 너비를 적절하게 설정하여 신경망을 설계하는게 좋다고 생각합니다.

**4장**

**(안정빈)**

**텍스트, 스크린샷, 소프트웨어, 운영 체제이(가) 표시된 사진

자동 생성된 설명텍스트, 스크린샷, 소프트웨어이(가) 표시된 사진

자동 생성된 설명**

**텍스트, 스크린샷, 소프트웨어, 멀티미디어 소프트웨어이(가) 표시된 사진

자동 생성된 설명**

**텍스트, 스크린샷, 폰트이(가) 표시된 사진

자동 생성된 설명**

**텍스트, 스크린샷, 소프트웨어, 멀티미디어 소프트웨어이(가) 표시된 사진

자동 생성된 설명텍스트, 스크린샷, 폰트, 소프트웨어이(가) 표시된 사진

자동 생성된 설명**

**텍스트, 스크린샷, 디스플레이, 소프트웨어이(가) 표시된 사진

자동 생성된 설명**

**텍스트, 스크린샷, 소프트웨어, 디스플레이이(가) 표시된 사진

자동 생성된 설명**

**텍스트, 스크린샷, 디스플레이, 소프트웨어이(가) 표시된 사진

자동 생성된 설명**

**텍스트, 스크린샷, 소프트웨어, 멀티미디어 소프트웨어이(가) 표시된 사진

자동 생성된 설명텍스트, 스크린샷, 소프트웨어, 멀티미디어 소프트웨어이(가) 표시된 사진

자동 생성된 설명**

**텍스트, 스크린샷, 소프트웨어, 멀티미디어 소프트웨어이(가) 표시된 사진

자동 생성된 설명**

**텍스트, 스크린샷, 디스플레이, 소프트웨어이(가) 표시된 사진

자동 생성된 설명**

**텍스트, 스크린샷, 소프트웨어, 멀티미디어 소프트웨어이(가) 표시된 사진

자동 생성된 설명**

**(김상옥)**

텍스트, 스크린샷, 번호, 폰트이(가) 표시된 사진

자동 생성된 설명

텍스트, 스크린샷, 소프트웨어, 번호이(가) 표시된 사진

자동 생성된 설명

텍스트, 스크린샷, 소프트웨어이(가) 표시된 사진

자동 생성된 설명

텍스트, 스크린샷, 번호, 소프트웨어이(가) 표시된 사진

자동 생성된 설명

텍스트, 스크린샷, 디스플레이, 직사각형이(가) 표시된 사진

자동 생성된 설명

텍스트, 스크린샷, 라인, 평행이(가) 표시된 사진

자동 생성된 설명

텍스트, 스크린샷, 번호, 평행이(가) 표시된 사진

자동 생성된 설명

텍스트, 스크린샷, 소프트웨어이(가) 표시된 사진

자동 생성된 설명텍스트, 스크린샷, 소프트웨어이(가) 표시된 사진

자동 생성된 설명 텍스트, 스크린샷, 소프트웨어, 번호이(가) 표시된 사진

자동 생성된 설명