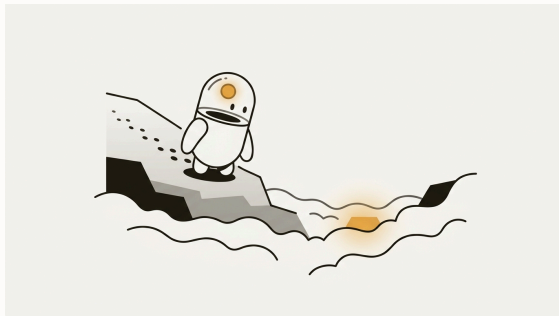




성능을 올리는 법

구조를 바꾸고 데이터를 손봐 점수를 올립니다

강의 영상에서 설명한 내용을 그대로 옮긴 교재입니다

정원석 지음 · Connect AI LAB

차례

이 챕터에서

-
- | | |
|---|---|
| 1 | 성능을 올리는 원리
백프로퍼게이션 · 렐루 · 드롭아웃 · 아담 |
|---|---|
-
- | | |
|---|--------------------------------|
| 2 | 구조를 바꿔 보면
층 하나, 둘, 셋 |
|---|--------------------------------|
-
- | | |
|---|----------------------------------|
| 3 | 순위표에 올리기
답안지를 만들어 제출한다 |
|---|----------------------------------|
-

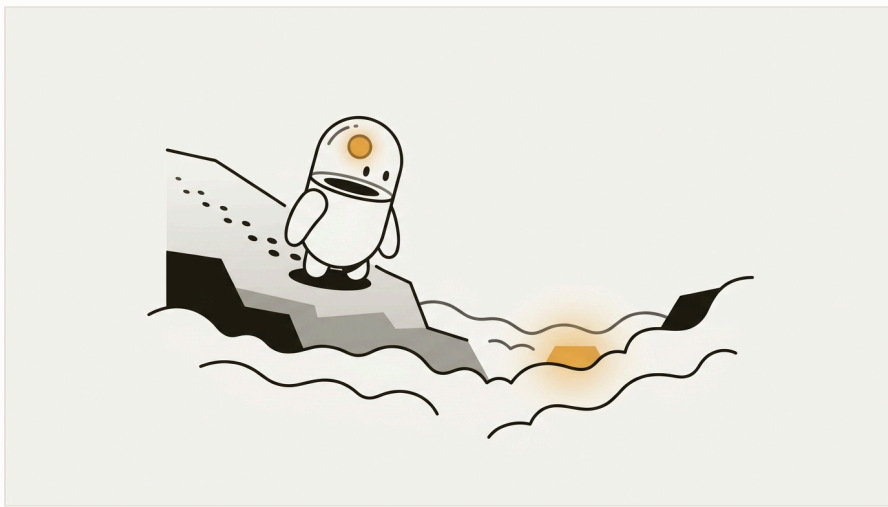
각 절 끝에 그 내용을 설명한 강의 영상이 붙어 있습니다. QR 을 찍으면 휴대폰에서도 열립니다.

5

성능을 올리는 원리

백프로퍼게이션 · 렐루 · 드롭아웃 · 아담

산을 내려간다



선을 긋고 거리를 재고 다시 긋는 일을 기계가 어떻게 할까요. 오차를 계산해서 **어느 쪽으로 고쳐야 하는지** 알려 줍니다. 이것을 **백프로퍼게이션**이라고 합니다.

그래프 위에서 로스가 거의 0 인 지점을 찾아 내려가는 일입니다. 선을 하나 긋고 계산해 보니 이쪽으로 가면 내려갈 것 같습니다. 가 봅니다. 또 계산합니다. 조금 더 내려갑니다.

그렇게 **더 이상 갈 수 없는 지점**까지 갑니다. 미분과 적분이 들어가지만 전부 알 필요는 없습니다. 거리를 재고 줄이는 방향을 찾는 것, 그 이름이 백프로퍼게이션이라는 것만 아시면 됩니다.

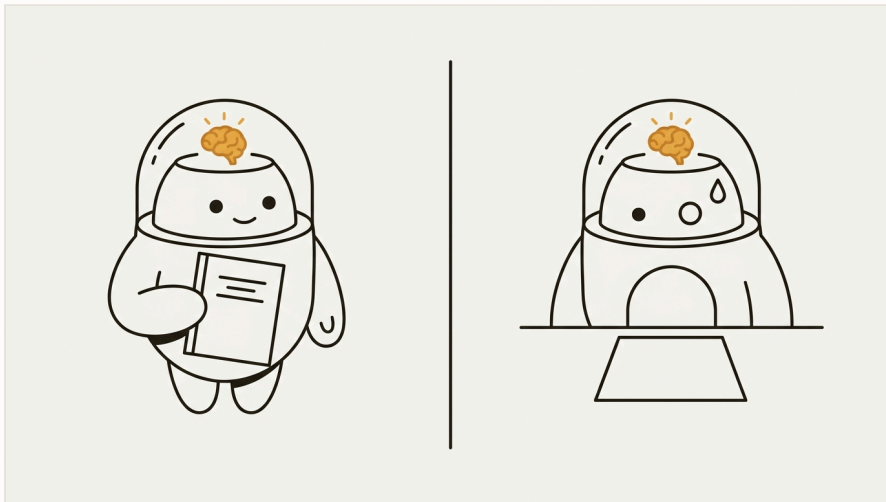
멈추는 것과 가는 것

내려가는 방법에도 여러 가지가 있습니다.

시그모이드를 쓰면 기울기가 완만해지는 구간에서 **그냥 멈춰 버립니다**. 차이가 크지 않으니 더 갈 이유를 못 찾는 것입니다. 로스가 1.51 쯤에서 멈춰 버린 적이 있습니다.

렐루는 터프합니다. 조금이라도 차이가 나면 **그냥 갑니다**. 그래서 더 내려갑니다.

많이 한다고 되지 않는다



그러면 무식하게 많이 훈련하면 될까요. 그렇지 않습니다.

데이터마다 다르지만 100번이 최적인 경우도 있고, 1,000번 하면 오히려 나빠지는 경우도 있습니다.

오버피팅 때문입니다. 연습한 데이터에서만 엄청나게 잘하고 실전에서는 못하는 상태입니다. 연습 문제는 90점인데 시험은 못 보는 학생과 같습니다.

반대로 너무 적게 훈련하면 **언더피팅**입니다. 연습에서도 못하고 실전에서도 못합니다.

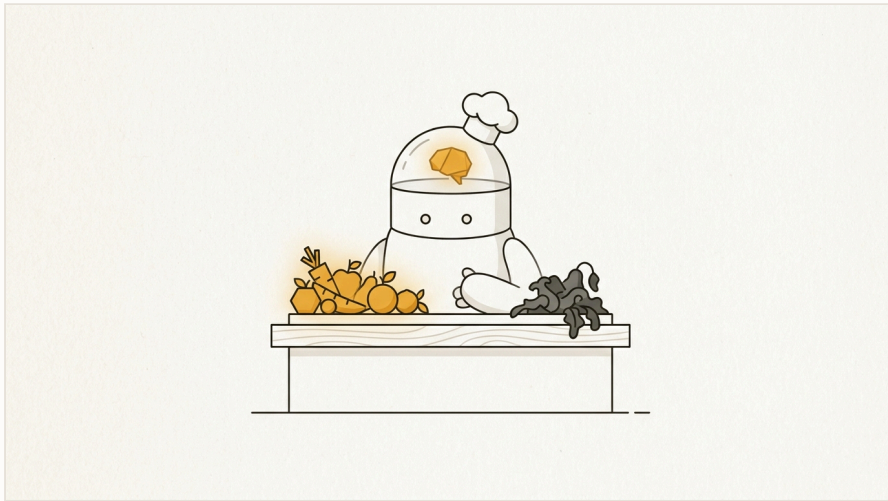
그 사이 어딘가, 적당히 일반화된 지점이 가장 좋습니다.

네 단어만 기억하세요

이름	하는 일
백프로퍼게이션	로스를 줄일 방향을 찾습니다
렐루	기울기가 완만해도 멈추지 않고 갑니다
드롭아웃	훈련할 때 일부를 일부러 꺼서 과적합을 막습니다
아담	내려갈 때의 보폭을 알맞게 잡아 줍니다

각각을 전부 공부하려면 아주 오래 걸립니다. **역할만 알아도 충분합니다.**

좋은 재료



아무리 좋은 요리사여도 재료가 나쁘면 좋은 요리를 만들 수 없습니다. 인공지능도 같습니다. **좋은 모델과 좋은 데이터가 함께 필요합니다.**

그래서 앞으로 모델을 만드는 사업만 잘 되는 게 아닙니다. 데이터를 모으고, 정답을 붙이고, 정제하는 일이 **어마어마하게 커집니다.**

글 데이터만이 아닙니다. 피지컬 AI 시뮬레이터가 모으는 데이터, 실제 로봇이 모으는 데이터가 그렇습니다.

왜냐하면 이것이 굉장히 귀찮은 작업이기 때문입니다. **사람은 남들이 하기 싫어하는 일을 할 때 돈을 벌니다.**

영상으로 보기

시작 — 성능을 높이는 원리

약 8분



youtu.be/QmK6zL9AVpU
휴대폰으로 찍으면 바로 열립니다

6

구조를 바꿔 보면

층 하나, 둘, 셋

층을 쌓아 봅니다

1강에서는 층 하나만 쌓았습니다. 이번에는 둘도 쌓고 셋도 쌓아 보고, 유명한 알고리즘도 넣어 보면서 무엇이 달라지는지 봅니다.

구조	시험 점수
층 한 줄	79.5%
층 두 줄	80.5%
층 세 줄	79.3%

두 줄에서 1% 올랐는데 세 줄에서는 오히려 떨어졌습니다. 복잡하게 만든다고 좋아지지 않습니다.

연습 점수와 시험 점수

점수를 볼 때는 두 개를 같이 봐야 합니다.

	점수
연습 점수	85.9%
시험 점수	79% 대

연습에서 90% 가 나오는데 시험에서 못 보는 것, 이것이 **오버피팅**입니다.

드롭아웃을 넣어 일부를 일부러 잊게 했더니 시험 점수가 **80.1%** 로 올라갔습니다.

호칭 하나로 81.3%

가장 크게 오른 것은 구조가 아니었습니다. 데이터였습니다.

이름을 보면 Mr, Mrs, Miss, Master 같은 호칭이 붙어 있습니다. 이 호칭 하나에 성별과 나이대와 결혼 여부가 함께 들어 있습니다.

모델은 그대로 두고 호칭 칸 하나만 추가했더니 81.3% 가 되었습니다.

같은 두뇌, 같은 도구. 어떤 칸을 주느냐로 갈립니다.

CNN과 트랜스포머

이미지 분석으로 유명한 CNN 을 넣어 봤습니다. 합성곱 신경망이라고 부릅니다. 결과는 **80.9%** 였습니다.

요즘 유행하는 **트랜스포머**도 넣어 봤습니다. GPT 의 T 가 바로 이것입니다. 문장 사이의 맥락을 잘 보는 구조이고, 「Attention is all you need」 라는 논문에서 나왔습니다. 결과는 **81.1%** 였습니다.

복잡하게 쓴다고 크게 달라지지 않았습니다. 전체를 놓고 보면 **79.5% 에서 81.3% 사이**였고, 가장 크게 올린 것은 호칭 한 칸이었습니다.

영상으로 보기

실습 — 두뇌 구조를 바꿔보면

약 20분 · 코랩



youtu.be/6cHID7j1XeQ
휴대폰으로 찍으면 바로 열립니다

7

순위표에 올리기

답안지를 만들어 제출한다

명단 두 개를 올립니다

파일이 두 개 있습니다. 가르칠 명단 891명과 시험 명단 418명입니다.

둘 다 내려받아서 코랩에 한 번에 선택해 업로드합니다. 가르칠 명단으로 배우고, 시험 명단으로 답을 만듭니다.

빈칸을 채우고, 쓸 열을 고르고, 재생을 누르면 답안지 파일이 자동으로 내려받아집니다.

올리면 순위가 나옵니다

받은 답안지를 대회 페이지에 올리면 바로 채점됩니다.

처음 올렸을 때 연습 점수가 61.6% 였습니다. 데이터를 잘못 골랐던 겁니다. 열을 바꾸고 다시 올려 **62.6%**, 구조를 바꿔 **80.6%** 로 올라갔습니다.

바꿀 것은 많습니다. 더 학습시켜 보고, 덜 학습시켜 보고, 요금과 탑승향을 넣어 보고, 다른 모델을 써 보세요.

90% 를 넘겨 보시기 바랍니다.

영상으로 보기

대회 — 답안지 만들어 순위표에 올리기

약 4분



youtu.be/A5eGTcxqznk
휴대폰으로 찍으면 바로 열립니다

맺음

1권을 마치며

여기까지 오셨다면 이제 **두뇌를 만들 줄 아는 사람**입니다. 함수가 무엇인지 알고, 기계가 그것을 스스로 만든다는 것을 알고, w 와 b 를 직접 꺼내 봤습니다.

2권에서는 **남이 만든 두뇌를 내려받아 씁니다**. 허깅페이스에서 고르고, 내 램에 맞는 크기를 정하고, 내 컴퓨터에서 돌립니다.

그다음 지식을 연결하고, 가중치를 바꾸고, 두 두뇌를 합칩니다. 그리고 귀와 입을 붙이면 자비스가 됩니다.

다음

챕터 3 · 두뇌 쇼핑

만들지 않습니다. 골라서 받아 수선향니다

챕터 3 읽기

강의 목록으로 돌아가려면 [여기](#)를 누르세요.