



8/4(수) 회의록

🕒 작성일시	@2021년 8월 3일 오후 7:46
👤 작성자	이하람
👥 참석자	
🕒 최종 편집일시	@2021년 8월 9일 오후 7:30
📌 회의 유형	일일 회의

👉 학습 내용 공유

1. 어제 과제 코드 리뷰 🔍

내일 화요일 과제를 코드 리뷰하는 것으로 하겠습니다.

월요일 과제) stop iteration 원인 : 0 입력 시 게임종료 로직에 문제가 있었다.

2. 오늘 과제 방법 토론타기 🤔

오늘 과제는 하신 분이 없어서 생략합니다.

3. 강의 내용 중 질문하기 🙋

Q. 필수 퀴즈 2강 - "어떠한" 행렬의 역행렬은 항상 계산할 수 있다?

⇒ "임의의" 행렬로 해석하면 될 것 같습니다.

Q. 경사하강법 - 다음 식에서 X^T 가 갑자기 나온 이유가 뭔가요?

$$\begin{aligned}\nabla_{\beta} \|y - X\beta\|_2^2 &= (\partial_{\beta_1} \|y - X\beta\|_2^2, \dots, \partial_{\beta_d} \|y - X\beta\|_2^2) \\ &= -\frac{2}{n} X^T (y - X\beta)\end{aligned}$$

$\|y - X\beta\|_2$ 대신 $\|y - X\beta\|_2^2$ 을 최소화하면 식이 좀 더 간단해진다

해설)

강의에서 사용된 RMSE Loss function $L = \left(\frac{\sum_i (y_i - \sum_j X_{ij} \omega_j)^2}{n} \right)^{1/2}$ 의 경사하강법을 살펴보면,
한 변수 ω_k 에 대한 편미분은 아래와 같이
$$\nabla_{\omega_k} L = -\frac{\sum_i X_{ik} (y_i - \sum_j X_{ij} \omega_j)}{\sqrt{n} (\sum (y_i - \sum_j X_{ij} \omega_j)^2)^{1/2}} = -\frac{\sum_i X_{ik} (y - X\omega)_i}{\sqrt{n} \|y - X\omega\|_2} = -\frac{(X_k)^\top (y - X\omega)}{\sqrt{n} \|y - X\omega\|_2}$$
가 되고,
따라서 L 의 gradient vector는 $\nabla_{\omega} L = (\nabla_{\omega_1} L, \dots, \nabla_{\omega_m} L) = -\frac{X^\top (y - X\omega)}{\sqrt{n} \|y - X\omega\|_2}$ 가 된다.
그러므로 점화식은 $\omega_n = \omega_{n-1} - \alpha \nabla_{\omega} L$ 로 나타나어진다.

Q. 활성화 함수와 비선형 함수의 차이가 뭔가요?

<활성 함수>

- 실수값을 입력으로 받아 다시 실수값을 반환해주는 비선형 함수
- 딥러닝에서 사용하는 활성화 함수

- 시그모이드(sigmoid) 함수

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

- tanh 함수

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- ReLU 함수

- 가장 많이 쓰임

$$\text{ReLU}(x) = \max\{0, x\}$$

- 잠재벡터(hidden vector)라고도 함
 - 선형모델로부터 나온 출력물을 비선형모델로 변환시킬 수 있음
 - 잠재벡터들을 뉴런(neuron)이라고도 부름
- softmax 와의 차이점
 - softmax 함수
 - 출력물의 모든 값을 고려해서 계산함
 - 활성화 함수
 - 해당 주소에 해당하는 출력값만 가지고 계산함
 - 벡터가 아닌 하나의 실수값을 input으로 받음

Q. 소프트맥스 함수는 활성화함수가 아닌건가요?

⇒ 소프트맥스 함수도 활성화함수다.

Q. "학습된다"는 것의 의미가 뭔가요?

⇒ 가중치(weight) 및 편향치(bias) 업데이트 과정을 학습이라고 표현해도 좋을 것 같아요!

Q. 추론 시에 소프트맥스 함수를 사용하지 않는다는 게 무슨 뜻인가요?

⇒ 학습이 끝난 모델의 경우에는 softmax 함수를 사용하지 않고 one-hot encoding

Q. 6강 확률론 21분 - 분류 문제에서 softmax를 계산하는 게 조건부 확률을 계산하는 것과 같다는게 무슨 의미인가요?

⇒ 입력데이터가 주어졌을 때, 그 데이터에 따라 softmax로 나온 결과가 각 카테고리별 확률 값이니까 결국에는 이 데이터에 대한 카테고리가 나올 조건부확률이다.

$P(Y|X)$ → 데이터 X가 카테고리 Y일 확률 → softmax와 상동

Q. 몬테카를로 샘플링이 뭔가요?

⇒ 원의 넓이를 계산하는 예시, 점을 무작위로 찍어서 계산? 이거 위키피디아에 나와있어요!

특징 : 가능한 모든 수를 시도하는 것이 전제로 들어감, 각 값이 맞는지 확률적으로

추후 자세한 내용을 알아올 것.

⇒ 약간 확률적인 완전탐색/브루트포스 같은 느낌?

(+) 독립추출이 보장되면, 대수의 법칙(law of large number)에 의해 수렴성 보장 = 수를 많이 뽑으면 정답에 가까워질 것이라는 뜻

Q. 연속형 확률변수의 한 지점에서의 밀도(density)는 그 자체로 확률값을 가진다?

연속형 확률변수는 확률밀도함수이므로, 각각의 점이 확률이 아니라 구간에 대한 적분이 확률이다.

밀도는 확률의 변화율이다.

점에서 적분은 불가능하기 때문에, 확률값이 존재하지 않는다?

Q. 9강 CNN 6분 째 수식이 무슨 뜻인가요?

'*' 는 convolution 연산을 뜻하고, kernel을 움직이면서 더한다는 것을 수학적으로 표현한 것이다.

'+' / '-' 의 뜻 : 함수의 좌우 반전!