

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/Проверка остаточных знаний/ПК-28

1. Алгоритм обратного распространения ошибки фактически

- a. ищет частные производные сложной функции
- b. вычисляет значение функции в точке
- c. тренирует нейронную сеть

Алгоритм обратного распространения ошибки фактически (Множественный выбор / Только один ответ)

2. В графе вычислений листья соответствуют...

- a. переменным
- b. искусственным нейронам
- c. функциям

В графе вычислений листья соответствуют (Множественный выбор / Только один ответ)

3. Есть два слоя нейросети, состоящие из 5 и 4 нейронов соответственно и связанные полностью. Сколько весов-связей существует между этими слоями?

- a. 20
- b. 9
- c. 5

Есть два слоя нейросети, состоящие из 5 и 4 нейронов соответственно и связанные полностью. (Множественный выбор / Только один ответ)

4. К выходящему из искусственного нейрона значению применяется

- a. функция активации
- b. взятие производной
- c. нормализация

К выходящему из искусственного нейрона значению применяется (Множественный выбор / Только один ответ)

5. Какой будет следующая точка при градиентном спуске для функции $y=x^2$ при начальной позиции $a_0=0$?

- a. 0
- b. 1
- c. -1

Какой будет следующая точка при градиентном спуске для функции (Множественный выбор / Только один ответ)

6. При каком значении аргумента достигается минимум функции $y=x^2-2x$?

- a. 0
- b. 1
- c. 2

При каком значении аргумента достигается минимум функции (Множественный выбор / Только один ответ)

7. Тренировочная выборка обязательно содержит

- a. Точные значения целевого признака
- b. Не менее 10 объектов
- c. Расписание тренировок

Тренировочная выборка обязательно содержит (Множественный выбор / Только один ответ)

8. У искусственного нейрона входное значение x имеет вес 2, вес-смещение равен 1, функция активации Relu. Какое значение выдаст нейрон, если входное значение было $x=-1$?

- a. 0
- b. -1
- c. 1

У искусственного нейрона входное значение x имеет вес 2, вес-смещение равен 1, функция активации Relu. Какое значение выдаст нейрон, если входное значение было $x=-1$? (Множественный выбор / Только один ответ)

9. Функция потерь для задачи регрессии и тренировочной выборки, состоящей из 100 объектов, состоит из ... слагаемых

- a. 100
- b. 10
- c. 1

Функция потерь для задачи регрессии и тренировочной выборки, состоящей из 100 объектов, состоит из ... слагаемых (Множественный выбор / Только один ответ)

10. Чему равна частная производная выражения $xу$ по переменной y ?

- a. x
- b. y
- c. 0

Чему равна частная производная выражения $xу$ по переменной y ? (Множественный выбор / Только один ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/Проверка остаточных знаний/ПК-29

1. Алгоритм стохастического градиентного спуска на каждой итерации выбирает ... слагаемые из функции потерь.

- a. одно
- b. все
- c. половину

Алгоритм стохастического градиентного спуска на каждой итерации выбирает ... слагаемые из функции потерь. (Множественный выбор / Только один ответ)

2. В идеале всю размеченную выборку нужно разбить на три части: тренировочную, тестовую и ...

- a. валидационную
- b. контрольную
- c. корректирующую

В идеале всю размеченную выборку нужно разбить на три части: тренировочную, тестовую и ... (Множественный выбор / Только один ответ)

3. В чем смысл преобразования softmax?

- a. Полученные числа можно интерпретировать как вероятности
- b. Ну какая же математическая формула без числа e ?
- c. Нужно для усложнения выражения функции потерь

В чем смысл преобразования softmax (Множественный выбор / Только один ответ)

4. Деактивация части нейронов на каждой итерации обучения нейросети называется...

- a. дропаут
- b. регуляризация
- c. нормализация по мини-батчам

Деактивация части нейронов на каждой итерации обучения нейросети называется... (Множественный выбор / Только один ответ)

5. Если нейронная сеть содержит вес w , то функция потерь при L1-регуляризации будет содержать слагаемое

- a. $C|w|$
- b. Cw^2
- c. Cw

Если нейронная сеть содержит вес w , то функция потерь при L1-регуляризации будет содержать слагаемое (Множественный выбор / Только один ответ)

6. Если нейронная сеть содержит вес w , то функция потерь при L2-регуляризации будет содержать слагаемое

- a. $C|w|$
- b. Cw^2
- c. Cw

Если нейронная сеть содержит вес w , то функция потерь при L2-регуляризации будет содержать слагаемое (Множественный выбор / Только один ответ)

7. При каком значении константы C регуляризация нейронной автоматически превратится в обычную тренировку?

- a. 0
- b. 1
- c. -1

При каком значении константы C регуляризация нейронной автоматически превратится в обычную тренировку? (Множественный выбор / Только один ответ)

8. Сколько каналов должен иметь фильтр, чтобы его можно было бы использовать при обработке изображения в формате RGB?

- a. любое натуральное число
- b. 3
- c. 1

Сколько каналов должен иметь фильтр, чтобы его можно было бы использовать при обработке изображения в формате RGB? (Множественный выбор / Только один ответ)

9. Чему равен результат пулинга 2x2 для изображения? 1234 5678
- a. 68
 - b. 48
 - c. 5678

Чему равен результат пулинга 2x2 для изображения? (Множественный выбор / Только один ответ)

10. Чему равен результат свёртки при применении фильтра 010 111 010 к изображению 111 101 111
- a. 4
 - b. 3
 - c. 0

Чему равен результат свёртки при применении фильтра (Множественный выбор / Только один ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/7. Распознавание изображений, свёрточные нейронные сети и фильтры

1. Было изображение 110 011 110 К нему сначала применили фильтр 11 а потом фильтр 0 1 Сколько в итоговом изображении пикселей равно 2?

Было изображение. Сколько в итоговом изображении пикселей равно 2? (Числовой ответ)

2. Дан фильтр: 010 111 010 и изображение: 0101010 111 1111 0101010 111 1111 0101010 Примените операцию свёртки при $stride_x=1$, $stride_y=2$ Сколько в результирующем изображении будет пикселей со значением 5?

Дан фильтр. Сколько в результирующем изображении будет пикселей со значением 5? (Числовой ответ)

3. Дан фильтр: 010 111 010 и изображение: 0101010 111 1111 0101010 111 1111 0101010 Сколько пикселей со значением 3 будет в результирующем изображении?

Дан фильтр. Сколько пикселей со значением 3 будет в результирующем изображении? (Числовой ответ)

4. Дан фильтр: 010 111 010 и изображение: 0101010 111 1111 0101010 111 1111 0101010 Примените операцию свёртки при $\text{stride}_x=2$, $\text{stride}_y=1$. Сколько столбцов пикселей будет иметь результирующее изображение?

Дан фильтр. Сколько столбцов пикселей будет иметь результирующее изображение? (Числовой ответ)

5. Дана нейронная сеть (не свёрточная), ее функция $FNN(x)=(w_1x+b_1)w_2+b_2$. Допустим, что её натренировали на некоторых данных и получили следующие значения весов: $w_1=1$, $b_1=1$, $w_2=1$, $b_2=-1$, и НС стала вычислять функцию: $FNN(x)=(1*x+1)*1-1=x$. Будем до-тренировывать эту НС на следующей ТВ:Теперь применим Transfer learning, установив разные значения шагов для каждого слоя. Для весов w_1, b_1 шаг равен $h_1=0.05$, а для остальных весов зададим шаг $h_2=0.1$ Найдите новые значения весов после одной итерации градиентного спуска. В ответ впишите новое значение веса b_2 .

Дана нейронная сеть (не свёрточная) впишите новое значение веса b_2 (Числовой ответ)

6. Дана нейронная сеть (не свёрточная), ее функция $FNN(x)=(w_1x+b_1)w_2+b_2$. Допустим, что её натренировали на некоторых данных и получили следующие значения весов: $w_1=1$, $b_1=1$, $w_2=1$, $b_2=-1$, и НС стала вычислять функцию: $FNN(x)=(1*x+1)*1-1=x$. Будем до-тренировывать эту НС на следующей ТВ:Теперь применим Transfer learning, установив разные значения шагов для каждого слоя. Для весов w_1, b_1 шаг равен $h_1=0.05$, а для остальных весов зададим шаг $h_2=0.1$ Найдите новые значения весов после одной итерации градиентного спуска. В ответ впишите новое значение веса w_1 .

Дана нейронная сеть (не свёрточная) впишите новое значение веса w_1 (Числовой ответ)

7. Дана нейронная сеть (не свёрточная), ее функция $FNN(x)=(w_1x+b_1)w_2+b_2$. Допустим, что её натренировали на некоторых данных и получили следующие значения весов: $w_1=1$, $b_1=1$, $w_2=1$, $b_2=-1$, и НС стала вычислять функцию: $FNN(x)=(1*x+1)*1-1=x$. Будем до-тренировывать эту НС на следующей ТВ:По смыслу Transfer learning мы заморозим веса w_1, b_1 , а для остальных весов зададим шаг $h=0.1$. Найдите новые значения весов после одной итерации градиентного спуска. В ответ впишите новое значение веса b_2 .

Дана нейронная сеть (не свёрточная). В ответ впишите новое значение веса b_2 (Числовой ответ)

8. Дана нейронная сеть (не свёрточная), ее функция $FNN(x)=(w_1x+b_1)w_2+b_2$. Допустим, что её натренировали на некоторых данных и получили следующие значения весов: $w_1=1$, $b_1=1$, $w_2=1$, $b_2=-1$, и НС стала вычислять функцию: $FNN(x)=(1*x+1)*1-1=x$. Будем до-тренировывать эту НС на следующей ТВ:По смыслу Transfer learning мы заморозим веса w_1, b_1 , а для остальных весов зададим шаг $h=0.1$ Найдите новые значения весов после одной итерации градиентного спуска. В ответ впишите новое значение веса w_2 .

Дана нейронная сеть (не свёрточная). В ответ впишите новое значение веса w_2 . (Числовой ответ)

- 9. Дано изображение и фильтр:**
- | | |
|-------------------------|-------------------------|
| 11011 | 101 |
| 10101 | 010 10101 |
| 101 11 111 00100 | 101 11 111 00100 |
- Нужно осуществить операцию деконволюции. Напоминаю, что сначала вычисляем размер деконволюции. Какой должно быть изображение, чтобы при применении фильтра 3×3 получилось бы изображение 5×5 ? А далее вычисляем все пиксели деконволюции. Сколько в результирующем изображении будет пикселей, равных 4?

Дано изображение и фильтр. Нужно осуществить операцию деконволюции (Числовой ответ)

- 10. Дано изображение 110 010 111 и фильтр 000 010 010. Добавим к изображению паддинг равный 1 и применим операцию свёртки. Сколько пикселей равно нулю в результирующем изображении?**

Дано изображение. Сколько пикселей равно нулю в результирующем изображении? (Числовой ответ)

- 11. Дано изображение 00000000 01110000 01111000 01111100 00111110 00011111 и фильтр 0 -1 0 -1 4 -1 0 -1 0. Сколько рядов пикселей будет в результирующем изображении?**

Дано изображение. Сколько рядов пикселей будет в результирующем изображении? (Числовой ответ)

- 12. Дано изображение 00000000 01110000 01111000 01111100 00111110 00011111 и фильтр 0 -1 0 -1 4 -1 0 -1 0. Сколько столбцов пикселей будет в результирующем изображении?**

Дано изображение. Сколько столбцов пикселей будет в результирующем изображении? (Числовой ответ)

- 13. Дано изображение 00000000 01110000 01111000 01111100 00111110 00011111 и фильтр 0 -1 0 -1 4 -1 0 -1 0. Сделайте операцию свёртки. Сколько пикселей со значением 1 будет в результирующем изображении?**

Ответ: 2

Сделайте операцию свёртки. Сколько пикселей со значением 1 будет в результирующем изображении? (Числовой ответ)

-
14. У изображения 2 канала: 1й канал 2й канал 010 110
111 011 К нему применяют двухканальный фильтр 1й канал 2й
канал 01 11 Какое сколько пикселей в результирующем
изображении равно 2?

У изображения 2 канала (Числовой ответ)

15. Чему равен пулинг и анпулинг? (Напоминание: применяется max-пулинг) Напишите в ответ значение ячейки q21.

Чему равен пулинг и анпулинг? Напишите в ответ значение ячейки q21 (Числовой ответ)

16. Чему равен пулинг и анпулинг? (Напоминание: применяется max-пулинг) Напишите в ответ значение ячейки q44.

Чему равен пулинг и анпулинг? Напишите в ответ значение ячейки q44. (Числовой ответ)

17. Что сделает фильтр $1/9 \ 1/9 \ 1/9 \ 1/9 \ 1/9 \ 1/9 \ 1/9 \ 1/9 \ 1/9$ с изображением 00000
01110 01010 01110 01010 01110 00000 Сколько там будет пикселей со
значением $1/3$?

Что сделает фильтр. Сколько там будет пикселей со значением $1/3$? (Числовой ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/9. Метрические задачи машинного обучения

1. При поиске минимума функции для задачи metric learning нужно минимизировать выражения с функцией max. Вот вам учебная функция: $L(w,b) = \max(2w+b-0.1, 0)$ Пусть $h=1$ и дана начальная точка: $w=2 \ b=1$ Чему будет равен вес w после одной итерации градиентного спуска?

При поиске минимума функции для задачи metric learning (Числовой ответ)

2. Пусть для объектов тройки (A,P,N) имеет место embedding во множество трехмерных векторов: $e(A)=(1,1,1)$ $e(P)=(1,1,0)$ $e(N)=(-1,-1,-1)$ Вычислить triplet loss для этой тройки при $\alpha=3$. $L(A,P,N) = \max(d(e(A),e(P)) - \alpha, 0)$

$d(e(A), e(N)) + \alpha, 0)$ В качестве метрики d использовать евклидову метрику для векторов.

Пусть для объектов тройки (A, P, N) имеет место embedding (Числовой ответ)

3. Пусть для объектов тройки (A, P, N) имеет место embedding во множество трехмерных векторов: $e(A) = (1, 0, 1)$ $e(P) = (0, 1, 0)$ $e(N) = (-1, 2, -1)$ Вычислить triplet loss для этой тройки при $\alpha = 2$. $L(A, P, N) = \max(d(e(A), e(P)) - d(e(A), e(N)) + \alpha, 0)$ В качестве метрики d использовать евклидову метрику для векторов. Все ответы округлить до 2-х знаков после запятой.

Пусть для объектов тройки (A, P, N) имеет место embedding (Числовой ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/10. Понятие о состязательных сетях (GAN)

1. Условимся считать положительные числа "правильными", а "отрицательные" - фейковыми. Посмотрим, как быстро генератор догадается генерировать только "правильные" числа. Сначала определимся с архитектурой G и D : Зададим начальные значения весов: $w=1, b=0, u=0, d=0$. Дискриминатор D возвращает вероятность, что вход x - фейковый, то есть отрицательный. Нужно сделать одну итерацию градиентного спуска для дискриминатора по выборке. Чему равен вес d после одной итерации градиентного спуска?

Чему равен вес d (Числовой ответ)

2. Условимся считать положительные числа "правильными", а "отрицательные" - фейковыми. Посмотрим, как быстро генератор догадается генерировать только "правильные" числа. Сначала определимся с архитектурой G и D : Зададим начальные значения весов: $w=1, b=0, u=0, d=0$. Дискриминатор D возвращает вероятность, что вход x - фейковый, то есть отрицательный. Нужно сделать одну итерацию градиентного спуска для дискриминатора по выборке. Чему равен вес генератора b после одной итерации градиентного спуска?

Чему равен вес генератора (Числовой ответ)

3. Условимся считать положительные числа "правильными", а "отрицательные" - фейковыми. Посмотрим, как быстро генератор догадается генерировать только "правильные" числа. Сначала определимся с архитектурой G и D : Зададим начальные значения весов: $w=1, b=0, u=0, d=0$. Дискриминатор D возвращает вероятность, что вход x - фейковый, то есть отрицательный. Нужно сделать одну итерацию градиентного спуска для дискриминатора по выборке. Чему равен вес генератора w после одной итерации градиентного спуска?

Чему равен вес генератора w (Числовой ответ)

4. Условимся считать положительные числа "правильными", а "отрицательные" - фейковыми. Посмотрим, как быстро генератор догадается генерировать только "правильные" числа. Сначала определимся с архитектурой G и D : Зададим начальные значения весов: $w=1, b=0, u=0, d=0$. Дискриминатор D возвращает вероятность, что вход x - фейковый, то есть отрицательный. Нужно сделать одну итерацию градиентного спуска для дискриминатора по выборке. Чему равен вес u после одной итерации градиентного спуска?

Чему равен вес u (Числовой ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/11. Обработка текстов

1. В языке первобытного племени 4 слова a, b, c, d . Дан один большой текст: $d\ c\ a\ b\ b\ b\ a\ c\ a\ d\ b\ a$. Будем тренировать модель `word2vec` по этому тексту. Для этого сформируйте тренировочную выборку (в виде таблицы), построенную с помощью 1-контекстов. Сколько строк с построенной таблице?

Будем тренировать модель `word2vec` (Числовой ответ)

2. В языке первобытного племени только 3 слова a, b, c . Дано три текста $A1: b\ a\ b\ a$, $A2: c\ c\ c\ a\ b\ c\ c\ c$, $A3: a\ a\ b\ a\ c\ a\ b\ a\ a$. Составьте для них таблицу вида: В ячейки таблицы запишите значение метрики TF. Какое значение будет лежать в ячейке $(A1, a)$?

В языке первобытного племени только 3 слова a, b, c (Числовой ответ)

3. В языке первобытного племени только 3 слова a, b, c . Дано три текста $A1: b\ a\ b\ a$, $A2: c\ c\ c\ a\ b\ c\ c\ c$, $A3: a\ a\ b\ a\ c\ a\ b\ a\ a$. Составьте для них таблицу вида: Заполните таблицу значениями метрики TF-IDF. В окошко ответа запишите значение TF-IDF для ячейки $(A1, c)$, ответ округлите до двух знаков после запятой.

В языке первобытного племени только 3 слова a, b, c . (Числовой ответ)

4. Задача на поиск синонимов. В языке гопников 4 слова a, b, c, d . Дан один большой текст: $d\ c\ a\ b\ b\ b\ a\ c\ a\ d\ b\ a\ a\ b\ a\ a\ c\ a\ b\ c\ b$. Нужно найти количество общих 1-контекстов у каждой пары слов. Пара с наибольшим количеством

одинаковых контекстов является парой слов-синонимов. Если вы всё сделали правильно, то парой с наибольшим количеством общих контекстов будут слова a , b . В ответ напишите количество их общих 1-контекстов.

Задача на поиск синонимов (Числовой ответ)

5. Задача на поиск синонимов. В языке гопников 4 слова a , b , c , d . Дан один большой текст: $d\ c\ a\ b\ b\ b\ a\ c\ a\ d\ b\ a\ a\ b\ a\ a\ c\ a\ b\ c\ b$ Нужно найти количество общих 1-контекстов у каждой пары слов. Пара с наибольшим количеством одинаковых контекстов является парой слов-синонимов. Посчитайте количество общих контекстов слов c и d .

Задача на поиск синонимов (копия) (Числовой ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/1. Поиск точки минимума функции с помощью градиентного спуска

1. Дана функция $f(x) = 3x^2 - 12x + 4$. Найдите у нее точку минимума (имеется в виду значение переменной x , при котором функция принимает минимальное значение). У функции ровно один минимум (как раз тот случай, когда локальный минимум совпадает с глобальным).

Вопрос 1 (Числовой ответ)

2. Функция $f(x) = 2x$. Теперь будем использовать вторую формулу численного дифференцирования при $\delta = 1$. Тогда $f'(0)$ приближенно равна...P.S. Ну как, применение более изоощренной формулы дало ли более точный ответ?

Вопрос 10 (Числовой ответ)

3. Вы искали минимум и строили график функции $f(x) = x(x^2 - 9)$. Из графика (при условии, что вы его правильно построили) следует, что "шарик" может скатиться как в точку минимума, так и заняться бесконечным спуском в бездонную пропасть слева. Давайте убедимся в том, что шарик из точки $a_0 = -1$ будет катиться по направлению к точке минимума, которую вы должны были найти на предыдущем шаге. Пусть $h = 0.1$, вычислите позицию шарика a_2 после двух шагов градиентного спуска и запишите ее в ответ. Действительно ли шарик приблизился к точке минимума?

Вопрос 11 (Числовой ответ)

4. И так, из графика функции $f(x)=x(x^2-9)$ следует, что "шарик" может скатиться как в точку минимума, так и заняться бесконечным спуском в бездонную пропасть слева. Давайте убедимся в том, что шарик из точки $a_0=-2$ будет катиться в бесконечную пропасть в левой области графика функции. Пусть $h=0.1$, вычислите позицию шарика a_2 после двух шагов градиентного спуска и запишите ее в ответ. Действительно ли шарик "совершил суицид" и начал лететь в пропасть?

Вопрос 12 (Числовой ответ)

5. Дана ФМП: Найдите ее частные производные (их 2 штуки). Приравняйте обе производные к нулю и решите полученную систему уравнений. Ее решение и будет точкой минимума. В ответ напишите значение переменной в точке минимума.

Вопрос 13 (Числовой ответ)

6. Дана функция: $f(x_1, x_2) = x_1^2 + 4x_1 + 2x_2^2 + 5x_2 + 3x_1x_2 + 6$, и у нее действительно есть точка минимума. Найдите ее и в ответ напишите координату точки минимума по второй переменной.

Вопрос 14 (Числовой ответ)

7. Дана функция, найти ее частные производные? Напишите в ответ, чему равно значение частной производной по второму аргументу в точке $(1, 0)$.

Вопрос 15 (Числовой ответ)

8. Задача на навыки дифференцирования. Дана функция, найти ее частные производные? Напишите в ответ, чему равно значение частной производной по первому аргументу в точке $(4, 2, 1)$.

Вопрос 16 (Числовой ответ)

9. Дана функция от двух аргументов $f(x_1, x_2) = x_1^2 + 10x_2^2$. Постройте ее график, будет видно, какая у этой функции точка минимума. Поверхность будет похожа на сплюснутый стаканчик. Пусть первая точка градиентного спуска равна $a_0 =$

(10, 1), шаг $h = 0.1$. Чему равна первая координата следующей позиции "шарика" $a_1 = (?, \dots)$? Впишите это значение в ответ.

Вопрос 17 (Числовой ответ)

10. Функция от двух аргументов $f(x_1, x_2) = x_1^2 + 10x_2^2$. Пусть первая точка градиентного спуска равна $a_0 = (10, 1)$, шаг $h = 0.1$. Чему равна первая координата позиции "шарика" после двух шагов градиентного спуска $a_2 = (?, \dots)$? Впишите это значение в ответ.

Вопрос 18 (Числовой ответ)

11. В задачах вы работали с функцией $f(x_1, x_2) = x_1^2 + 10x_2^2$ и нашли позицию a_2 шарика" после двух шагов градиентного спуска. А чему равна вторая координата точки a_2 ? Напишите это значение.

Вопрос 19 (Числовой ответ)

12. Дана функция $f(x) = x(x^2 - 9)$. Где у нее точка минимума? Напишите значение аргумента x , при котором у функции наблюдается локальный минимум. А ещё лучше постройте ее график – пригодится на следующем этапе. В ответе написать число с двумя знаками после запятой.

Вопрос 2 (Числовой ответ)

13. Пусть $f(x) = x^2$, $a_0 = 1$, $h = 1$, тогда следующей точкой спуска будет $a_1 = -1$, а потом будет $a_2 = 1$ - то есть градиентный спуск вернётся к первоначальной позиции (проверьте это сами!). Иными словами, градиентный спуск «заиклится» и на следующих шагах мы уже никогда не приблизимся к точке минимума $x = 0$. Мы будем лишь колебаться между точками -1 и 1 . Что делать? А нужно применить одну из модификаций градиентного спуска, связанной с изменением шага h . Итак, возьмем формулу $h = h_0 (1 - n/T)$ для $h_0 = 1$, $T = 4$. Это означает, что возникнет следующая последовательность шагов: $1, 0.75, 0.5, 0.25, 0$. Кроме того, после пятого шага спуска шаг станет равным 0 и процесс самопроизвольно остановится. В общем, у вас есть 5 шагов, чтобы подойти к точке минимума. Давайте вычислять позиции "шарика" после каждого шага. Поскольку вначале шаг равен 1 , то (как и ранее) мы получим $a_1 = -1$. А вот дальше начинаются различия. Найдите позиции шарика a_2 , a_3 , не забывая, что в первом случае вы должны использовать шаг $h = 0.75$, а во втором случае $h = 0.5$. В ответ запишите точку a_3 .

Вопрос 20 (Числовой ответ)

14. Возьмём функцию $f(x) = x^2$, $a_0 = 1$, $h = 1$. Ранее мы выяснили, что "классический" градиентный спуск будет лишь колебаться между точками -1 и 1 . Что делать? А нужно применить одну из модификаций градиентного спуска, связанной с изменением шага h . На этот раз возьмем другую формулу: $h = h_0 e^{-p/T}$ для $h_0 = 1$, $T = 1$. Это означает, что возникнет такая последовательность уменьшающихся шагов h : $1, 0.37, 0.14 \dots$. Давайте вычислять позиции «шарика» после каждого шага. Поскольку вначале шаг равен 1 , то (как и ранее) мы получим $a_1 = -1$. А вот дальше начинаются различия. Найдите позиции шарика a_2 , a_3 , не забывайте, что в первом случае вы должны использовать шаг $h=0.37$, а во втором случае $h=0.14$. В ответ запишите точку a_3 .

Вопрос 21 (Числовой ответ)

15. Дана функция $f(x_1, x_2) = x_1^2 + 10x_2^2$. Ранее мы выяснили, что при $a_0 = (10, 1)$, $h = 0,1$ значения второй координаты нашего "шарика" будут прыгать между 1 и -1 . При этом значения по первой координате очень медленно подбираются к нулю. Метод Adagrad способен настроить разную величину шага по каждой из координат. В нашем случае нам необходимо, чтобы шаг по первой координате увеличивался, а по второй координате - уменьшался. Сделаем два шага этого метода и получим точки a_1 , a_2 . В дальнейших расчетах полагаем в ответе напишите, чему равна первая координата точки a_2 .

Вопрос 22 (Числовой ответ)

16. Дана функция $f(x_1, x_2) = x_1^2 + 10x_2^2$. Ранее мы выяснили, что при $a_0 = (10, 1)$, $h = 0,1$ значения второй координаты нашего "шарика" будут прыгать между 1 и -1 . При этом значения по первой координате очень медленно подбираются к нулю. Как ускорить спуск по первой координате и одновременно избежать заикливания по второй координате? Посмотрим, что сделает метод Adam! В дальнейших расчетах полагаем в ответе напишите, чему равна первая координата точки a_2 .

Вопрос 23 (Числовой ответ)

17. Дана парабола $f(x)=x^2$. Представьте себе, что наш "шарик" находится в точке с координатой $a_0=2$. Мы запускаем процедуру градиентного спуска из точки a_0 с шагом $h=0.1$. Найдите позицию шарика на следующем шаге a_1 и позицию a_2 через два шага спуска. В ответ написать только значение a_2 . Ну как? Действительно ли шарик катится туда, куда нужно?

Вопрос 3 (Числовой ответ)

18. Со школы мы знаем, что чем круче график функции идёт вверх, тем больше значение производной (ибо производная - это скорость изменения функции). Естественно, крутизна графика влияет на ускорение нашего "шарика" из процедуры градиентного спуска. Вот вам теперь функция $f(x)=x^4$. Представьте себе, что наш "шарик" находится в точке с координатой $a_0=2$. Мы запускаем

процедуру градиентного спуска из точки a_0 , а с шагом $h=0.1$. Найдите позицию шарика на следующем шаге a_1 . Ну как: он прошел большее расстояние, чем в предыдущей задаче (где мы брали функцию $f(x)=x^2$)? В ответе будет отрицательное число. Это означает, что шарик настолько разогнался, что проскочил точку минимума $x=0$. Взрыв градиента, однако!

- 1.2
- 1.2

Вопрос 4 (Числовой ответ)

19. Вы теперь думаете, что шарик по графику функции $f(x)=x^4$ всегда катится быстрее, чем по графику функции $f(x)=x^2$? А вот и нет! Можете найти все такие начальные позиции a_0 , что градиентный спуск по функции $f(x)=x^2$ приведет к не меньшему перемещению шарика, чем по функции $f(x)=x^4$? В ответе напишите максимально возможное значение для такой начальной позиции a_0 .

- 0.71
- 0.71

Вопрос 5 (Числовой ответ)

20. Дана функция $f(x)=x^4$, и начальная позиция равна $a_0=1$. Чему должен быть равен шаг обучения h , чтобы на следующей итерации градиентного спуска мы оказались бы в точке $a_1=-1$? (и напишите это значение h в ответ) P.S. На досуге подумайте еще вот о чём: а если в этой задаче взять функцию $f(x)=x^6$, то ответ будет больше (меньше) и почему? P.S.S. Кто догадается, куда из точки $a_1=-1$ при найденном значении шага обучения h попадёт шарик на втором шаге? Иными словами, чему равно a_2 ?

Вопрос 6 (Числовой ответ)

21. Устроим градиентный спуск по функции $f(x)=1/x$. Из ее графика понятно, что при больших положительных x у функции имеется плато, на котором градиентный спуск практически останавливается (затухание градиента). Чему равна позиция шарика a_0 , если известно, что его следующая позиция a_1 находится на расстоянии всего-то от позиции 10^{-6} . Шаг обучения равен $h=0.01$.

Вопрос 7 (Числовой ответ)

22. Дана функция $f(x)=1/x$. Из ее графика понятно, что при малых положительных x у функции имеется крутой участок, на котором возможен взрыв градиента. Чему равна позиция шарика a_0 , если известно, что его следующая позиция a_1 "улетела" на расстояние 100 от позиции a_0 ? Шаг обучения равен $h=0.01$.

Вопрос 8 (Числовой ответ)

23. Займёмся численным дифференцированием. Итак, $f(x)=2x$. Конечно, мы можем явно вычислить ее производную: $f'(x)=2x \ln 2$. Но представим себе, что у нас нет доступа к выражению для функции f , и мы можем лишь запрашивать значение функции в той или иной точке. Как найти производную нашей функции в точке 0? Будем использовать первую формулу численного дифференцирования при $\delta=1$. Тогда $f'(0)$ приближенно равна... P.S. Кстати, $f'(0)=20 \ln 2 \approx 0.693$

Вопрос 9 (Числовой ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/2. Графы вычислений и алгоритм обратного распространения ошибки

1. Задание на применение школьной формулы дифференцирования сложной функции. Чему равна производная функции $f(x) = e \sin x$? В ответ напишите значение .

Вопрос 1 (Числовой ответ)

2. Чему равна производная функции? В ответ напишите значение

Вопрос 2 (Числовой ответ)

3. Задача на применение формулы про дифференцирование сложной функции. Пусть Найдите частные производные сложной функции по переменным Не забудьте потом вхождения букв заменить на их выражения В ответ напишите значение (т. е. в выражение для частной производной подставьте значения P.S. В принципе, правильный ответ можно получить и без знания формулы дифференцирования сложной функции. Нужно сразу буквы заменить на их определения и получить выражение для функции А потом найти ее частные производные с помощью школьной формулы дифференцирования частного. Так что можете применить этот метод и сравнить полученные ответы. Разумеется, ответ при применении обоих методов должен быть одинаковым.

Вопрос 3 (Числовой ответ)

4. Задача на применение формулы про дифференцирование сложной функции. Пусть Найдите частные производные сложной функции по переменным Не забудьте потом вхождения букв заменить на их выражения через буквы В ответ напишите значение P.S. Правильный ответ можно получить и без знания

формулы дифференцирования сложной функции. Нужно сразу буквы заменить на их определения и получить выражение для функции. А потом найти ее частные производные с помощью школьной формулы дифференцирования произведения. Так что можете применить этот метод и сравнить полученные ответы. Разумеется, ответ при применении обоих методов должен быть одинаковым.

Вопрос 4 (Числовой ответ)

5. Задача на применение формулы про дифференцирование сложной функции. Пусть T есть фактически функция f зависит от одной переменной x . Найдите производную сложной функции по ее единственной переменной x . Не забудьте потом вхождения букв заменить на их выражения. В ответ напишите значение

Вопрос 5 (Числовой ответ)

6. Можете построить граф вычислений, если указана последовательность операций, которые применяются к переменным. Итак, у нас будет одна переменная x (это будет единственный лист графа). Далее идут вершины $a := 2x$, $b := 3x$, $c := a+b$, $d := ab$, $f := c/d$, где f - корень графа (он соответствует всей функции $f(x)$). Последовательно найдите частные производные $\partial f/\partial c$, $\partial f/\partial d$, $\partial f/\partial a$, $\partial f/\partial b$, $\partial f/\partial x$. Собственно, последняя частная производная в списке и будет окончательным ответом. В окошко ответа запишите значение $\partial f/\partial x(1)$. P.S. Свои вычисления можно проверить "в лоб". Очевидно, что можно раскрутить выражение для f и получить выражение для f через букву x . Должно получиться $f(x) = (2x + 3x)/6x$ (понятно, откуда шестёрка появилась?). После этого выражение можно продифференцировать по школьным правилам, подставить $x=1$ и получить ответ.

Вопрос 6 (Числовой ответ)

7. Можете построить граф вычислений, если указана последовательность операций, которые применяются к переменным. Итак, у нас будет теперь две переменных x_1 , x_2 (значит, в графе будет 2 листа). Далее идут вершины $a := x_1 x_2$, $b := x_1 + x_2$, $c := ab$, $d := a+b$, $f := c^2 + d^2$, где f - корень графа (он соответствует всей функции $f(x_1, x_2)$). Последовательно найдите частные производные. Собственно, последние две частных производных в списке и будут окончательным ответом. В окошко ответа запишите значение. P.S. Свои вычисления можно проверить "в лоб". Очевидно, что можно раскрутить выражение для f и получить выражение для f через буквы x_1 , x_2 . Должно получиться $f(x_1, x_2) = (x_1^2 x_2 + x_1 x_2^2)^2 + (x_1 x_2 + x_1 + x_2)^2$. После этого выражение можно продифференцировать по школьным правилам, подставить $x_1 = 1$, $x_2 = 1$ и получить ответ.

Вопрос 7 (Числовой ответ)

8. Производные суперпозиции нескольких сигмоид имеют достаточно предсказуемый внешний вид: это произведения выражений, содержащих те же сигмоиды. Так, например, выражение для производной одиночной сигмоид содержит множитель. Выражение для производной суперпозиции двух сигмоид содержит множители. Не надо быть пророком, чтобы догадаться, что производная суперпозиции n сигмоид будет содержать n множителей (и понятно каких множителей). Поскольку все множители в выражении для производных сигмоид положительны и меньше 1, то справедливы неравенства: (данные неравенства получаются из выражений для производной, в котором отброшена часть множителей. Поскольку отбрасываются множители меньшие 1, то равенство превратится в неравенство со знаком $<$). Теперь мы можем воспользоваться определением сигмоиды и оценить производные в точке $x=0$. Получаем: Напишите, какое число получается в неравенстве для производной.

Вопрос 8 (Числовой ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/3. Тренировка нейросети для задачи регрессии

1. Представьте себе такой искусственный нейрон: ему на вход подается число x , затем оно домножается на вес связи, внутри нейрона прибавляется вес-смещение и после этого применяется функция активации Relu. Нарисуйте схему такого нейрона. Допустим, что в наш нейрон подали вход $x=4$ и получили ответ 1, потом подали вход $x=2$ и получили ответ 3. Чему равны веса? В окошко ответа напишите выход нейрона при $x=0$.

Вопрос 1 (Числовой ответ)

2. Снова берём нейросеть и тренировочную выборку. Для человека не составит труда восстановить зависимость между x и y . Ну да, $y=x+1$. То есть оптимальные значения весов сети. Посмотрим, сумеет ли нейронная сеть обнаружить эти значения в процессе тренировки. Для этого нужно составить функцию потерь $L(w)$, и начать процедуру градиентного спуска. Для старта спуска возьмем причем мы предполагаем, что первая координата позиции "шарика" соответствует весу w_1 а вторая координата соответствует весу w_2 . Давайте сделаем два шага градиентного спуска и получим позицию. В окошко ответа запишите первую координату (которая соответствует весу w_1) позиции. Ну как? Стала ли позиция "шарика" ближе к точке истинной минимума функции потерь $(1,1)$?

Вопрос 10 (Числовой ответ)

3. И снова нейросеть и тренировочная выборка. Для человека не составит труда восстановить зависимость между x и y . И так, $y=x+1$, т. е. оптимальные значения весов сети. Посмотрим, сумеет ли нейронная сеть обнаружить эти значения в процессе тренировки. Для этого нужно составить функцию потерь $L(w)$, и начать процедуру градиентного спуска. В этой задаче мы будем применять алгоритм стохастического градиентного спуска. Для старта спуска возьмем причем мы

предполагаем, что первая координата позиции "шарика" соответствует весу a вторая координата соответствует весу b . Давайте сделаем два шага градиентного спуска и получим позицию. Поскольку спуск у нас стохастический, то давайте условимся, что на первой итерации был выбран первый объект тренировочной выборки, а на второй итерации спуска был выбран второй объект тренировочной выборки. В окошко ответа запишите первую координату (которая соответствует весу a) позиции. Ну как? Стала ли позиция "шарика" ближе к точке истинной минимума функции потерь $(1,1)$?

Вопрос 11 (Числовой ответ)

4. Вот вам старая-добрая нейросеть И тренировочная выборка. Для человека не составит труда восстановить зависимость между x и y . И так, $y = -x$, т. е. оптимальные значения весов сети. Посмотрим, сумеет ли нейронная сеть обнаружить эти значения в процессе тренировки. Сейчас мы будем минимизировать функцию потерь $L(w)$ с помощью стохастического градиентного спуска по мини-батчам. Размер батча будет равен 2. Для старта спуска возьмем $w_1 = 1$ и $w_2 = 1$, причем мы предполагаем, что первая координата позиции "шарика" соответствует весу a вторая координата соответствует весу b . Давайте сделаем два шага стохастического градиентного спуска по мини-батчам и получим позицию. Мы предполагаем, что первой итерации будет выбран батч из 1го и 2го объекта тренировочной выборки, а на второй итерации батч будет состоять из 2го и 3го объекта выборки. В окошко ответа запишите вторую координату (которая соответствует весу b) позиции. Ну как? Стала ли позиция "шарика" ближе к точке истинной минимума функции потерь $(0,-1)$?

Вопрос 12 (Числовой ответ)

5. Представьте себе такой искусственный нейрон: ему на вход подается число x , затем оно домножается на вес связи w , внутри нейрона прибавляется вес-смещение b и после этого применяется функция активации Relu. Допустим, что в наш нейрон подали вход $x=4$ и получили ответ 1, потом подали вход $x=2$ и получили ответ 3. Чему равен выход нейрона при входе $x=10$?

Вопрос 2 (Числовой ответ)

6. Теперь представьте себе такой искусственный нейрон: ему на вход подается число x , затем оно домножается на вес связи w , внутри нейрона прибавляется вес-смещение b и после этого применяется функция активации сигмоида σ . Нарисуйте схему такого нейрона. Допустим, что в наш нейрон подали вход $x=1$ и получили ответ 0.5, потом подали вход $x=2$ и снова получили ответ 0.5. Чему равны веса w и b ? В окошко ответа напишите значение w .

Вопрос 3 (Числовой ответ)

7. Теперь представьте себе такой искусственный нейрон: у него два входа x_1 и x_2 , и эти входы нейрон преобразует в значение z . Нарисуйте схему такого нейрона. Как

видно, у него вообще нет функции активации. В наш нейрон поступают не одиночные числа, а пары чисел. Допустим, что наш нейрон выдал следующие ответы: . Чему равны веса ? В окошко ответа напишите значение .

Ответ: 3

Вопрос 4 (Числовой ответ)

8. Представьте себе такую многослойную нейронную сеть: у нее входной слой размерности 4, далее идут внутренние полносвязные слои, состоящие соответственно из 10, 20, 5 нейронов. Вся сеть заканчивается выходным слоем из одного нейрона. Сколько в этой сети весов-связей и весов-смещений? Напишите в ответ общее количество весов всех типов.

Вопрос 5 (Числовой ответ)

9. Представьте себе такую многослойную нейронную сеть: у нее входной слой размерности 1; далее идут два внутренних полносвязных слоя с неизвестным числом нейронов, причем количество нейронов на первом внутреннем слое больше числа нейронов на втором внутреннем слое; вся сеть заканчивается выходным слоем из одного нейрона. Известно, что во всей сети число весов-смещений равно 10, а весов-связей 29. Найдите, чему равно количество нейронов на втором внутреннем слое, и запишите это в окошко ответа.

Вопрос 6 (Числовой ответ)

10. Ниже приведена нейросеть с известными значениями весов. Все функции активации - это Relu. Чему равен выход нейросети, если $x=2$?

Вопрос 7 (Числовой ответ)

11. Возьмем нейросеть из прошлой задачи. Чему может быть равен x , если известно, что нейросеть выдала ответ 12?

Вопрос 8 (Числовой ответ)

12. Вот вам простенькая нейросеть и тренировочная выборка. Для человека не составит труда восстановить зависимость между x и y . Ну да, $y=-x$. То есть оптимальные значения весов сети. Посмотрим, сумеет ли нейронная сеть обнаружить эти значения в процессе тренировки. Для этого нужно составить функцию потерь $L(w)$, и начать процедуру градиентного спуска (будем

применять классический градиентный спуск без всяких наворотов). Для старта спуска возьмем причем мы предполагаем, что первая координата позиции "шарика" соответствует весу а вторая координата соответствует весу Давайте сделаем два шага градиентного спуска и получим позицию В окошко ответа запишите вторую координату, которая соответствует весу позиции Ну как? Стала ли позиция "шарика" ближе к точке истинной минимума функции потерь (0,-1)?

Вопрос 9 (Числовой ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/4. Технологии тренировки глубоких нейронных сетей

1. Дана таблица. У нее 3 нецелевых признака, которые нужно нормализовать. Мы будем работать с формулой нормализации $x_i := x_i - a - b \frac{x_i - a}{b - a}$, где а и b - соответственно минимальное и максимальное значение в столбце. Впишите в поле, чему будет равно значение признака "Вес" у третьего объекта после нормализации.

Вопрос 1 (Числовой ответ)

2. Вот вам простенькая нейросеть И тренировочная выборка Для человека не составит труда восстановить зависимость между x и y. Ну да, $y = -x$. То есть оптимальные значения весов сети Посмотрим, сумеет ли нейронная сеть обнаружить эти значения в процессе тренировки. В этой задаче мы будем применять L1-регуляризацию, а именно: припишем к "обычной" функции потерь где в качестве константы С мы возьмем $C = 1$. После этого начинаем обычную процедуру градиентного спуска. Для старта спуска возьмем причем мы предполагаем, что первая координата позиции "шарика" соответствует весу а вторая координата соответствует весу Давайте сделаем два шага градиентного спуска и получим позицию В окошко ответа запишите вторую координату (которая соответствует весу) позиции

Вопрос 10 (Числовой ответ)

3. Дропаут будем постигать на следующей нейросети: В этой сети часть весов уже натренирована и не подлежит изменению в процессе градиентного спуска. Тренировочная выборка такая: Начальные значения всех весов равны 0. Для человека очевидно, что искомая зависимость здесь $Y = X + 1$. Допустим, что на первой итерации был уничтожен верхний нейрон. Чему равен вес $w_2 w_2$ после одного шага градиентного спуска? Шаг спуска $h = 0.1$

Вопрос 11 (Числовой ответ)

4. Дропаут будем постигать на следующей нейросети: В этой сети часть весов уже натренирована и не подлежит изменению в процессе градиентного спуска. Тренировочная выборка такая: Начальные значения всех весов равны 0. Для человека очевидно, что искомая зависимость здесь $Y=X+1$. Допустим, что на первой итерации был уничтожен верхний нейрон. А чему равен вес w_{02} после первой итерации градиентного спуска?

Вопрос 12 (Числовой ответ)

5. Теперь освежим действие дропаута на входном слое следующей нейросети: Тренировочная выборка очевидна для человека. Еще бы: $Y=X_1 \cdot X_2$. Но посмотрим, к какой зависимости придет нейросеть с применением дропаута на входном слое. Допустим, что вначале был удален вход x_2 . Проведите одну итерацию градиентного спуска и напишите, чему будет равен вес w_{11}

Вопрос 13 (Числовой ответ)

6. Теперь освежим действие дропаута на входном слое следующей нейросети: Тренировочная выборка очевидна для человека. Еще бы: $Y=X_1 \cdot X_2$. Но посмотрим, к какой зависимости придет нейросеть с применением дропаута на входном слое. Допустим, что вначале был удален вход x_2 . Проведите одну итерацию градиентного спуска. Итак, мы провели одну итерацию градиентного спуска для нейросети. Теперь предположим, что на второй итерации был удален вход x_1 . Чему будет равен вес w_{00} после второго шага спуска. Не забудьте, что этот вес был изменен на первой итерации спуска, его текущее значение не равно 0.

Вопрос 14 (Числовой ответ)

7. Дана таблица. У нее 3 нецелевых признака, которые нужно нормализовать. Мы будем работать с формулой нормализации $x_i = \frac{x_i - a}{b - a}$, где a и b - соответственно минимальное и максимальное значение в столбце. Впишите в поле, чему будет равно значение признака "Группа крови" у второго объекта после нормализации. Каково это иметь группу крови в виде дробного числа?))))

Вопрос 2 (Числовой ответ)

8. Дана таблица. У нее 3 нецелевых признака, которые нужно нормализовать. Мы будем работать с формулой нормализации $x_i = \frac{x_i - a}{b - a}$, где a и b - соответственно минимальное и максимальное значение в столбце. Теперь допустим, что нейросеть как-то натренировалась на нормализованных данных. Теперь нам нужно пропустить через сеть новый объект: Но вначале его нужно нормализовать с помощью параметров, которые были вычислены по тренировочной выборке. Напишите в поле, во что превратится число 90 (признак "Вес") после нормализации?

Вопрос 3 (Числовой ответ)

9. Теперь мы будем нормализовывать те же самые данные с помощью формулы где \bar{x} - среднее значение столбца, S - величина отклонения в столбце. Напишите в поле, чему будет равно значение признака "Вес" у третьего объекта после нормализации. Вы не забыли, что мы используем формулу нормализации с отклонением и средним значением?

Вопрос 4 (Числовой ответ)

10. Теперь мы будем нормализовывать те же самые данные с помощью формулы где \bar{x} - среднее значение столбца, S - величина отклонения в столбце. Напишите в окошечко, чему будет равно значение признака "Пол" у второго объекта после нормализации. Теперь он у нас небинарная личность!)))

Вопрос 5 (Числовой ответ)

11. Теперь мы будем нормализовывать те же самые данные с помощью формулы где \bar{x} - среднее значение столбца, S - величина отклонения в столбце. Допустим, что нейросеть как-то натренировалась на нормализованных данных. Теперь нам нужно пропустить через сеть новый объект: Но вначале его нужно нормализовать с помощью параметров, которые были вычислены по тренировочной выборке. Напишите в окошечко, во что превратится число 90 (признак "Вес") после нормализации.

Вопрос 6 (Числовой ответ)

12. И снова нейросеть и тренировочная выборка Для человека не составит труда восстановить зависимость между x и y . Ну да, $y=x+1$. То есть оптимальные значения весов сети. Посмотрим, сумеет ли нейронная сеть обнаружить эти значения в процессе тренировки. В этой задаче мы будем применять регуляризацию, а именно: припишем к "обычной" функции потерь где C в качестве константы C мы возьмем $C=1$. После этого начинаем обычную процедуру градиентного спуска. Для старта спуска возьмем причем мы предполагаем, что первая координата позиции "шарика" соответствует весу w_1 а вторая координата соответствует весу w_2 . Давайте сделаем два шага градиентного спуска и получим позицию (w_1, w_2) в окошко ответа запишите первую координату, которая соответствует весу позиции. Ну как? Стала ли позиция "шарика" ближе к точке истинной минимума функции потерь $(1,1)$? Быстрее или медленнее теперь (то есть и с использованием регуляризации) веса стремятся к оптимальным значениям? Сравните свой ответ с ответом аналогичной задачи, которая решалась без использования регуляризации. Спойлер: на самом деле веса медленнее сходятся к оптимальным значениям. Мешает регуляризация, ибо она негодует, когда веса нейросети увеличиваются (по модулю).

Вопрос 7 (Числовой ответ)

13. Вот вам простенькая нейросеть И тренировочная выборка Для человека не составит труда восстановить зависимость между x и y . Ну да, $y=-x$. То есть оптимальные значения весов сети Посмотрим, сумеет ли нейронная сеть обнаружить эти значения в процессе тренировки. В этой задаче мы будем применять регуляризацию, а именно: припишем к "обычной" функции потерь где в качестве константы C мы возьмем $C=1$. После этого начинаем обычную процедуру градиентного спуска. Для старта спуска возьмем причем мы предполагаем, что первая координата позиции "шарика" соответствует весу а вторая координата соответствует весу Давайте сделаем два шага градиентного спуска и получим позицию В окошко ответа запишите вторую координату (которая соответствует весу) позиции Ну как? Стала ли позиция "шарика" ближе к точке истинной минимума функции потерь $(0,-1)$? Спойлер: на самом деле веса медленнее сходятся к оптимальным значениям. Мешает регуляризация, ибо она негодует, когда веса нейросети увеличиваются (по модулю).

Вопрос 8 (Числовой ответ)

14. И снова нейросеть и тренировочная выборка Для человека не составит труда восстановить зависимость между x и y . Ну да, $y=x+1$. То есть оптимальные значения весов сети Посмотрим, сумеет ли нейронная сеть обнаружить эти значения в процессе тренировки. В этой задаче мы будем применять L1-регуляризацию, а именно: припишем к "обычной" функции потерь где в качестве константы C мы возьмем $C=1$. После этого начинаем обычную процедуру градиентного спуска. Для старта спуска возьмем причем мы предполагаем, что первая координата позиции "шарика" соответствует весу а вторая координата соответствует весу Давайте сделаем два шага градиентного спуска и получим позицию В окошко ответа запишите первую координату (которая соответствует весу) позиции Совет. Если возникли сложности с вычислением производной выражения, содержащей модуль, то посмотрите подсказку: Как дифференцировать выражения с модулем? Все очень просто. Например, вам нужно найти частную производную функции потерь по переменной Функция потерь при L1-регуляризации имеет вид: (мы работаем со значением $C=1$). Классическую функцию потерь вы дифференцируете по обычным правилам, а производная выражения равна функции Например, если вам нужно подсчитать значение производной функции в точке Для этого вы просто пользуетесь формулой для производной модуль-функции, приведённой выше. И получаете ответ -1 .

Вопрос 9 (Числовой ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/5. Задача классификации для нейронных сетей

1. Чему равно значение функции потерь для данной нейросети-классификатора?

Вопрос 1 (Числовой ответ)

-
2. И так, у нас была нейросеть-классификатор: А теперь мы рассмотрим другой классификатор (с другими показателями точности): Восстановите значение верхней-правой ячейки, если известно, что функция потерь данного классификатора равна функции потерь классификатора с предыдущего шага. Ответ округлите до двух знаков после запятой. Решение задачи потребует навыков решения уравнений "с логарифмом".

Вопрос 2 (Числовой ответ)

3. У нас был такой классификатор: А теперь у нас появился второй классификатор, про который известно, что значение его функции потерь 1.5 раза меньше, чем у первого классификатора. Восстановите значение пропущенной верхней правой ячейки (и запишите ее в ответ, округлив до двух знаков после запятой).

Вопрос 3 (Числовой ответ)

4. Рассмотрим тренировочную выборку и будем тренировать нейронную сеть со следующей архитектурой: Человек в принципе понимает смысл зависимости между x и y в тренировочной выборке: все ненулевые числа принадлежат классу 1, а число "ноль" принадлежит классу с меткой 0. Но сможет ли до этого догадаться нейросеть? Составьте функцию потерь по заданной тренировочной выборке. Там должно быть три слагаемых (по количеству объектов из тренировочной выборки). Для простоты дифференцирования функцию потерь преобразуйте к виду (здесь надо пользоваться свойствами логарифмов и почленным делением дробей): Теперь найдите частные производные по всем переменным функции потерь, пользуясь школьным правилами дифференцирования. Прекрасно! А теперь зададим параметры градиентного спуска. Будем считать, что изначально все веса равны 0, шаг обучения равен $h=0.1$ Найдите новые значения весов после одного шага градиентного спуска. В окошко запишите новое значение веса

Вопрос 4 (Числовой ответ)

5. Рассмотрим тренировочную выборку и будем тренировать нейронную сеть со следующей архитектурой: Человек в принципе понимает смысл зависимости между x и y в тренировочной выборке: все ненулевые числа принадлежат классу 1, а число "ноль" принадлежит классу с меткой 0. Но сможет ли до этого догадаться нейросеть? Составьте функцию потерь по заданной тренировочной выборке. Там должно быть три слагаемых (по количеству объектов из тренировочной выборки). Для простоты дифференцирования функцию потерь преобразуйте к виду (здесь надо пользоваться свойствами логарифмов и почленным делением дробей): Теперь найдите частные производные по всем переменным функции потерь, пользуясь школьным правилами дифференцирования. А теперь зададим параметры градиентного спуска. Будем считать, что изначально все веса равны 0, шаг обучения равен $h=0.1$ Найдите новые значения весов после одного шага градиентного спуска. В окошко запишите новое значение веса Спойлер: Посмотрев на тренировку нейросети, можно сделать далеко идущие

выводы.Нейросеть решила схалтурить и оставить веса нулями , а тренировать только веса-смещения Это приведет к тому, что возникнет так называемый дурацкий классификатор. Такой классификатор для каждого объекта выдает одни и те же вероятности, основываясь на процентном соотношении представителей классов в тренировочной выборке.

Вопрос 5 (Числовой ответ)

- 6. Рассмотрим тренировочную выборкуи будем тренировать нейронную сеть со следующей архитектурой:Составьте функцию потерь по заданной тренировочной выборке. Там должно быть два слагаемых (по количеству объектов из тренировочной выборки). Для простоты дифференцирования функцию потерь преобразуйте к виду (тут надо пользоваться свойствами логарифмов и почленным делением дробей):Теперь найдите частные производные по всем переменным функции потерь (пользуясь школьным правилами дифференцирования).А теперь зададим параметры градиентного спуска. Будем считать, что изначально все веса равны 0, шаг обучения равен $h=0.1$ Найдите новые значения весов после одного шага градиентного спуска. В окошко запишите новое значение веса или веса (если вы все сделали правильно, то значения этих весов будут одинаковыми).**

Вопрос 6 (Числовой ответ)

- 7. Рассмотрим тренировочную выборкуи будем тренировать нейронную сеть со следующей архитектурой:Составьте функцию потерь по заданной тренировочной выборке. Там должно быть два слагаемых (по количеству объектов из тренировочной выборки). Для простоты дифференцирования функцию потерь преобразуйте к виду (тут надо пользоваться свойствами логарифмов и почленным делением дробей):Теперь найдите частные производные по всем переменным функции потерь (пользуясь школьным правилами дифференцирования).А теперь зададим параметры градиентного спуска. Будем считать, что изначально все веса равны 0, шаг обучения равен $h=0.1$ Найдите новые значения весов после одного шага градиентного спуска. Напишите новое значение веса (если вы все сделали правильно, то новый вес будет иметь противоположное значение). Подумайте, насколько логичен такой результат тренировки нейросети.**

Вопрос 7 (Числовой ответ)

- 8. «Два дебила – это сила»Рассмотрим два дурацких классификатора K1 и K2.Дурацкий классификатор K1 был натренирован по тренировочной выборке, содержащей 5 объектов класса 0 и 5 объектов класса 1.Дурацкий классификатор K2 был натренирован по тренировочной выборке, содержащей 10 объектов класса 0 и 10 объектов класса 1.Оба классификатора выдают одинаковые ответы (они для любого объекта говорят, что он с вероятностью 0.5 принадлежит классу 0 и с вероятностью 0.5 принадлежит классу 1). Однако значения функции потерь на тренировочной выборке у них различны (вот это поворот!).Найдите значение функции потерь для каждой тренировочной выборки и в ответ запишите максимальное значение.Спойлер. Если вы всё сделали правильно, то функция потерь выше у тренировочной выборки классификатора K2. Это означает, что K2 "более глуп", чем K1. Дело в том, что**

K2 имел в своем распоряжении большую выборку, чем K1, но ничего кроме дурацкой стратегии классификации из нее не извлёк. Поэтому и функция потерь у него больше.

Вопрос 8 (Числовой ответ)

9. Допустим, что наша тренировочная выборка состоит из 4-х объектов. Три объекта A1, A2, A3 в реальности принадлежат классу 0, а один объект A4 в реальности принадлежит классу 1. Как будет работать дурацкий классификатор K1 на такой выборке? Подсчитайте его функцию потерь. А теперь рассмотрим другой (не-дурацкий) классификатор K2. Он для объектов A1, A2, A3 выдал вероятность принадлежности классу 0 равную 0.6. Чему равна вероятность принадлежности объекта A4 к классу 1 по мнению классификатора K2, если известно, что функция потерь K2 на всей тренировочной выборке численно равна функции потерь классификатора K1?

Вопрос 9 (Числовой ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/6. Контроль качества натренированных нейронных сетей

1. Допустим, что у нас была тренировочная выборка. По ней была натренирована некоторая нейросеть, которая на самом деле вычисляет функцию: Как видно, для всех объектов тренировочной выборки предсказанное нейросетью значение совпадает с истинным значением целевого признака. Но тут оказывается, что валидационная выборка (которая, естественно, была недоступна на этапе тренировки) представляет собой таблицу: По данной валидационной выборке подсчитайте MAE. Эта величина уж точно будет ненулевой!

Вопрос 1 (Числовой ответ)

2. А вот вам задача из жизни на взаимодействие заказчика и исполнителей. Заказчик нанял команду дата-саентистов и снабдил их следующей тренировочной выборкой: Заказчик сказал, что оплатит работу, если значение MAE на валидационной выборке будет не меньше 0.05. Валидационную выборку заказчик, естественно, держит в секрете. Дата-саентисты не сильно напрягались и натренировали нейросеть, которая фактически выдает константное значение (значение a пока неизвестно). Валидационная выборка была и модель действительно показала на ней MAE=0.05. Заказчик остался доволен и еще сильнее уверовал в силу машинного обучения. Напишите в окно ответа максимально возможно значение параметра a , при котором могла происходить описанная выше жизненная ситуация.

Вопрос 2 (Числовой ответ)

3. Займемся валидационной выборкой для задачи классификации. На этот раз команда дата-саентистов тренирует нейросеть, которая по результатам анализа принимает решение, болен (класс 1) или здоров (класс 0) пациент. Нейросеть была натренирована и ей дали некоторую валидационную выборку (ВВ), состоящую из 100 объектов. Известно, что двух истинных больных нейросеть классифицировала как здоровых, у троих истинных больных были найдены признаки болезни, а пятеро истинно здоровых людей по совету нейросети загремели в больницу. Подсчитайте precision и recall класса 1 для указанной ВВ. Напишите в ответ значение recall.

Вопрос 3 (Числовой ответ)

4. Команда дата-саентистов тренирует нейросеть, которая по результатам анализа принимает решение, болен (класс 1) или здоров (класс 0) пациент. Нейросеть была натренирована и ей дали некоторую валидационную выборку (ВВ), состоящую из 100 объектов. Известно, что двух истинных больных нейросеть классифицировала как здоровых, у троих истинных больных были найдены признаки болезни, а пятеро истинно здоровых людей по совету нейросети загремели в больницу. На эту нейросеть стали жаловаться следующие группы людей: 1) истинно больные люди, которых нейросеть приняла за здоровых 2) здоровые люди, которые из-за нейросети очутились в больнице (нейросеть работала не в военкомате, поэтому жалобы такого типа действительно имели место быть). Дата-саентисты подкрутили немного параметры сети и для контроля подали на вход нейросети такую же валидационную выборку (ВВ), что и в предыдущей задаче. Оказалось, что теперь уже четверо истинно больных были классифицированы как больные, и precision для класса 1 оказалась равна 0.4 Как видно, precision и recall возросли по сравнению с прошлой задачей. Но следует ли из этого, стало меньше здоровых людей, которые классифицируются нейросетью как больные? Напишите в ответе количество здоровых людей, отнесенных нейросетью к классу больных.

Вопрос 4 (Числовой ответ)

5. Допустим, что по некоторой валидационной выборке мы получили precision=0.7 recall=0.8 Найдите значение агрегированной характеристики F и запишите найденное значение в ответ. P.S. Как изменился бы ответ, если поменять местами значения: precision=0.8 recall=0.7 ? Сможете ответить на этот вопрос, не проводя вообще никаких вычислений?

Вопрос 5 (Числовой ответ)

6. Значение величины F одинаково для обоих случаев: 1) precision=0.7 recall=0.8 2) precision=0.8 recall=0.7 В задачах, где precision и recall имеют разную значимость (например, в задаче о выявлении больных: признать больного здоровым - это более серьезная ошибка, чем признать больным здорового) применяют характеристику F_{β} и мы также поступим. Найдите для $\beta=2$ значение F_{β} для precision=0.7 recall=0.8, а потом для случая precision=0.8 recall=0.7. У вас получится в итоге два числа. В ответ напишите максимальное из них.

Вопрос 6 (Числовой ответ)

7. Пусть теперь Подсчитайте значение для $\text{precision}=0.7$ $\text{recall}=0.8$, а потом для случая $\text{precision}=0.8$ $\text{recall}=0.7$. У вас получится в итоге два числа. В ответ как и раньше напишите максимальное из них.

Вопрос 7 (Числовой ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/8. Автокодировщики

1. Между кодирующей и декодирующей частями автокодировщика находится...
- бутылочное горлышко
 - свёрточный слой
 - полносвязный слой

Вопрос 1 (Множественный выбор / Только один ответ)

2. В задаче подавления шумов автокодировщик тренируется на парах изображений (A, B) , где A - это...
- зашумленное изображение
 - "чистое" изображение
 - матрица шумов

Вопрос 2 (Множественный выбор / Только один ответ)

3. Пусть при использовании автокодировщика были получены embedding-и объектов A , B : $e(A)=(1,2)$, $e(B)=(3,4)$. Тогда промежуточный объект, соответствующий середине между векторами $e(A)$, $e(B)$ имеет вторую координату равную...

Вопрос 3 (Числовой ответ)

4. Пусть при использовании автокодировщика были получены embedding-и фотографий улыбающихся людей A , B : $e(A)=(1,2)$, $e(B)=(2,2)$ и фотографий грустных людей $e(C)=(5,0)$ $e(D)=(4,2)$. Найдите вектор, соединяющий центр группы грустных людей и с центром группы веселых людей. Чему равна его первая координата?

Вопрос 4 (Числовой ответ)

5. Пусть при использовании автокодировщика были получены embedding-и фотографий улыбающихся людей A , B : $e(A)=(1,2)$, $e(B)=(2,2)$ и фотографий

грустных людей $e(C)=(5,0)$ $e(D)=(4,2)$. Имеем фотографию грустного человека с embedding-ом $e(F)=(3,3)$. При наложении улыбки на эту фотографию к вектору $e(F)$ был прибавлен вектор W , соединяющий центры множества точек грустных и веселых людей. Чему равна вторая координата вектора $e(F)+W$?

Вопрос 5 (Числовой ответ)

По умолчанию для Введение в нейронные сети (09.04.01, 2021)/12. Рекуррентные нейронные сети

1. Какие типы данных логичнее обрабатывать с помощью рекуррентных сетей?
- изображения
 - тексты
 - аудио
 - видео
 - численные векторы

Вопрос 1 (Множественный выбор)

2. Дана следующая архитектура рекуррентной нейронной сети: Пусть длина истории равна $t=2$. Чему равна функция, вычисляемая сетью?
- $FNN(x_2)=f(x_2w_2+b_2+vf(x_1w_1+b_1+v))$
 - $FNN(x_2)=f(x_2w+b+vf(x_1w+b+v))$
 - $FNN(x_2)=f(x_2w+b)$

Вопрос 2 (Множественный выбор / Только один ответ)

3. Дана следующая архитектура рекуррентной нейронной сети: Какое выражение вычисляется на втором выходе сети?
- $FNN(x_2)=f(x_2w_2+b_2+vf(x_1w_1+b_1+v))$
 - $FNN(x_2)=f(x_2w+b+vf(x_1w+b+v))$
 - $FNN(x_2)=f(x_2w+b)$

Вопрос 3 (Множественный выбор / Только один ответ)
