

- > Подготовка к НЭ по Анализу данных.
- > Продвинутый уровень
- > Тренировочный вариант 2 (из реального экзамена)

Оставшееся время 2:59:13

Вопрос 12

Пока нет
ответа

Балл: 0,63

В выборке 600 объектов положительного класса и 400 объектов отрицательного класса.

Алгоритм бинарной классификации на всех объектах выборки предсказывает одну и ту же вероятность попасть в положительный класс, то есть

$a(x) = \beta$, где β - некоторое число из отрезка $[0; 1]$

Чему равен ROC-AUC алгоритма, посчитанный на этой выборке? Ответ округлите до десятых.

Ответ:

Вопрос 13

Пока нет
ответа

Балл: 0,63

Дана выборка для задачи регрессии:

x	y
3	10
2	5
5	15
6	14

Эта задача решается при помощи решающего дерева.

Разбиения в дереве ищутся по условию $x > \theta$, где θ - такой порог, для которого значение Information Gain (IG) максимально.

$IG = H(R) - \frac{|R_l|}{|R|}H(R_l) - \frac{|R_r|}{|R|}H(R_r)$, где $H(A)$ - несмещенная дисперсия целевой переменной в вершине A.

Посчитайте Information Gain для разбиения по условию $x > 5$.

Ответ округлите до десятых.

Ответ:

Вопрос 14Пока нет
ответа

Балл: 0,63

Алгоритм word2vec кодирования текстов закодировал словосочетания "бутерброд с колбасой", "бутерброд без колбасы", "бутерброд с сыром" векторами $a = (1, 2, 3, 2, 4)$, $b = (0.5, 2, 2.5, 2, 4)$ и $c = (5, 2, 2, 3, 4)$ соответственно.

Известно, что операции над полученными word2vec-векторами сохраняют смысл текстов, то есть, например,

(вектор слова "король") + (вектор слова "женщина") = (вектор слова "королева").

Найдите вектор словосочетания "бутерброд без сыра". В ответ запишите квадрат его длины (евклидовой нормы).

Ответ: **Вопрос 15**Пока нет
ответа

Балл: 0,63

В проекте по машинному обучению для обработки данных используются три скрипта. Первый скрипт обрабатывает 40% данных, второй - свои 35% данных, третий - оставшиеся 25% данных. Вероятность некорректной работы первого скрипта составляет 0.04, второго - 0.06, а третьего - 0.03.

В результате работы скриптов на некотором объекте в данных обнаружена ошибка.

Какова вероятность того, что ошибка допущена вторым скриптом?

Ответ округлите до сотых.

Ответ:

Предыдущий
элемент курса

Тренировочный
вариант 1 UPDATE

Перейти на...

Следующий
элемент курса

Cheat Sheet 1



Служба
поддержки сайта

Вы зашли под именем
Царахова Милена
Викторовна (Выход)