

Семинар 6-7: Удержание клиентов

Задача 1

Бандерлог начинает все определения со слов «это доля правильных ответов»:

1. ассигасу — это доля правильных ответов...
2. точность (precision) — это доля правильных ответов...
3. полнота (recall) — это доля правильных ответов...
4. TPR — это доля правильных ответов...

Закончите определения Бандерлога так, чтобы они были, хм, правильными.

1. accuracy = $\frac{TP+TN}{TP+FP+FN+TN}$
2. precision = $\frac{TP}{TP+FP}$
3. recall = $\frac{TP}{TP+FN}$
4. TPR = $\frac{TP}{TP+FN}$

Задача 2

Бандерлог из Лога¹ ведёт блог, любит считать логарифмы и оценивать логистические регрессии. С помощью нового алгоритма Бандерлог решил задачу классификации по трём наблюдениям и получил $b_i = \hat{p}(y_i = 1|x_i)$.

y_i	b_i
1	0.7
-1	0.2
-1	0.3

1. Постройте ROC-кривую.
2. Найдите площадь под ROC-кривой и индекс Джини.
3. Постройте PR-кривую (кривая точность-полнота).
4. Найдите площадь под PR-кривой.
5. Как по-английски будет «бревно»?

Задача 3

Задача 3

¹деревня в Кадуйском районе Вологодской области

Задача 4

1. Ещё задачи

Задача 5

Задача 6

Задача 7