

Дано: задача классификации.

$X^\ell = \{x_1, \dots, x_\ell\}$ — выборка;

$y_i = y(x_i) \in \{0, 1\}$, $i = 1, \dots, \ell$ — известные бинарные ответы.

$b: X \rightarrow \mathbb{R}$ — алгоритм, оценивающий принадлежность x к классу 1.

Вопрос:

Как измерить качество $b(x)$ на выборке X^ℓ ?

Как правило, классификатор имеет вид

$$a(x) = [b(x) > t].$$

- $b(x)$ — оценка принадлежности к классу 1
- t — порог классификации

Линейный классификатор:

$$a(x) = [\langle w, x \rangle > 0].$$

[здесь идет картинка с разделяющей прямой; скалярное произведение оценивает расстояние; если объект близко к ней, то мы не уверены, если далеко, то уверены в ответе]

Метод k ближайших соседей:

$$a(x) = \left[\sum_{i=1}^k [y^{(i)} = 1] > k/2 \right].$$

[здесь идет визуализация тоже: если среди соседей объекта все относятся к одному классу, то он уверен в классификации и выдает высокую оценку; если же среди соседей встречаются оба класса, то оценка понижается]

Пример: кредитный скоринг.

- нужно предсказать, вернет ли клиент кредит;
- сортируем клиентов по оценке вероятности возврата $b(x)$;
- банк получает ранжированный список;
- порог выбирается в зависимости от стратегии банка;
- порог может многократно пересматриваться.

Пример: задача определения самолетов противника.

- $b(x)$ оценивает вероятность того, что самолет принадлежит противнику;
- классификатор $a(x) = [b(x) > 0.9]$;
- $\text{precision} = 0.2$, $\text{recall} = 0.7$;
- как понять, в чем проблема — неправильном пороге или плохой функции $b(x)$?

- 1 Отсортируем объекты по возрастанию оценки $b(x)$:

$$b(x_{(1)}) \leq \dots \leq b(x_{(\ell)}).$$

- 2 Переберем все пороги классификации, начав с максимального:

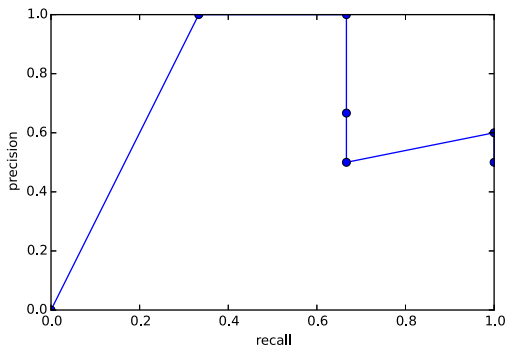
$$t_{\ell} = b(x_{(\ell)}),$$

...

$$t_1 = b(x_{(1)}),$$

$$t_0 = b(x_{(1)}) - \varepsilon.$$

- 3 Для каждого порога посчитаем точность и полноту.
- 4 Нанесем соответствующую точку в осях «полнота-точность».
- 5 Соединим точки, получив Precision-Recall-кривую.



$b(x)$	0.14	0.23	0.39	0.52	0.73	0.90
y	0	1	0	0	1	1

Свойства:

- Левая точка: всегда $(0, 0)$ (все объекты относим к классу 0);
- Правая точка: $(1, \ell_+/\ell)$, ℓ_+ — число объектов класса 1 в выборке;
- Если выборка идеально разделима, то кривая пройдет через точку $(1, 1)$;
- Чем больше площадь под кривой, тем лучше.

AUC-PRC (Area Under Precision-Recall curve) — мера качества для $b(x)$.

ROC — «reciever operating characteristic».

- по оси X: False Positive Rate, доля ошибочных положительных классификаций:

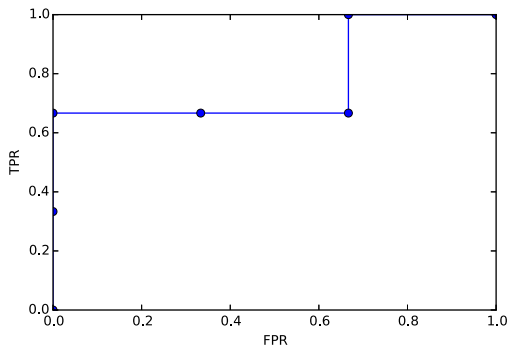
$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}.$$

1 – FPR называется *специфичностью* алгоритма.

- по оси Y: True Positive Rate, доля правильных положительных классификаций:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}.$$

TPR называется *чувствительностью* алгоритма.



$b(x)$	0.14	0.23	0.39	0.52	0.73	0.90
y	0	1	0	0	1	1

Свойства:

- Левая точка: всегда $(0, 0)$ (все объекты относим к классу 0);
- Правая точка: всегда $(1, 1)$ (все объекты относим к классу 1);
- Если выборка идеально разделима, то кривая пройдет через точку $(1, 0)$;
- Площадь меняется от $1/2$ до 1;
- Чем больше площадь под кривой, тем лучше.

AUC-ROC (Area Under ROC-curve) — мера качества для $b(x)$.

ROC-кривая:

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}, \quad \text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}.$$

- Метрики качества нормируются на размеры классов, ROC-кривая не изменится при перемене соотношения классов.
- Интерпретация: AUC-ROC равен вероятности того, что случайно взятый объект класса 1 получит оценку выше, чем случайно взятый объект класса 0.
- Имеет проблемы при сильном дисбалансе классов.

PR-кривая:

$$\text{precision} = \frac{TP}{TP + FP}, \quad \text{recall} = \frac{TP}{TP + FN}.$$

- Точность нормируется на число положительных прогнозов, изменится при перемене соотношения классов.
- Максимально возможная площадь под PR-кривой зависит от соотношения классов.
- Хорошо подходит для измерения качества при сильном дисбалансе классов.

- 100 объектов класса 1;
- 1.000.000 объектов класса 0;
- Ранжирование: 50.000 объектов класса 0, затем 100 объектов класса 1, затем все остальные объекты класса 0;

Метрики качества:

- AUC-ROC: 0.95;
- AUC-PRC: 0.001.

Почему так получается?

- Выберем порог, при котором первые 50.095 объектов относятся к классу 1;
- $TPR = 0.95$, $FPR = 0.05$;
- $precision = 0.0019$, $recall = 0.95$.

- Работать с оценками принадлежности может быть полезнее, чем с бинарными ответами
- Две основные метрики качества: AUC-PRC и AUC-ROC
- AUC-ROC не зависит от соотношения классов