Линейные модели классификации - II.

Recap

Модель линейной классификации

Классификация

- $\mathbb{Y} = \{-1, +1\}$
- -1 — отрицательный класс
- $+1$ — положительный класс
- $a(x)$ должен возвращать одно из двух чисел

Линейная регрессия

$$a(x) = \boxed{w_0 + \sum_{j=1}^{d} w_j x^j}$$

Вещественное число!

Линейный классификатор

$$a(x) = \operatorname{sign}\left(w_0 + \sum_{j=1}^{d} w_j x^j\right)$$

Линейный классификатор

$$a(x) = \operatorname{sign}\left(w_0 + \sum_{j=1}^{d} w_j x^j\right)$$

Свободный коэффициент

Веса

Признаки

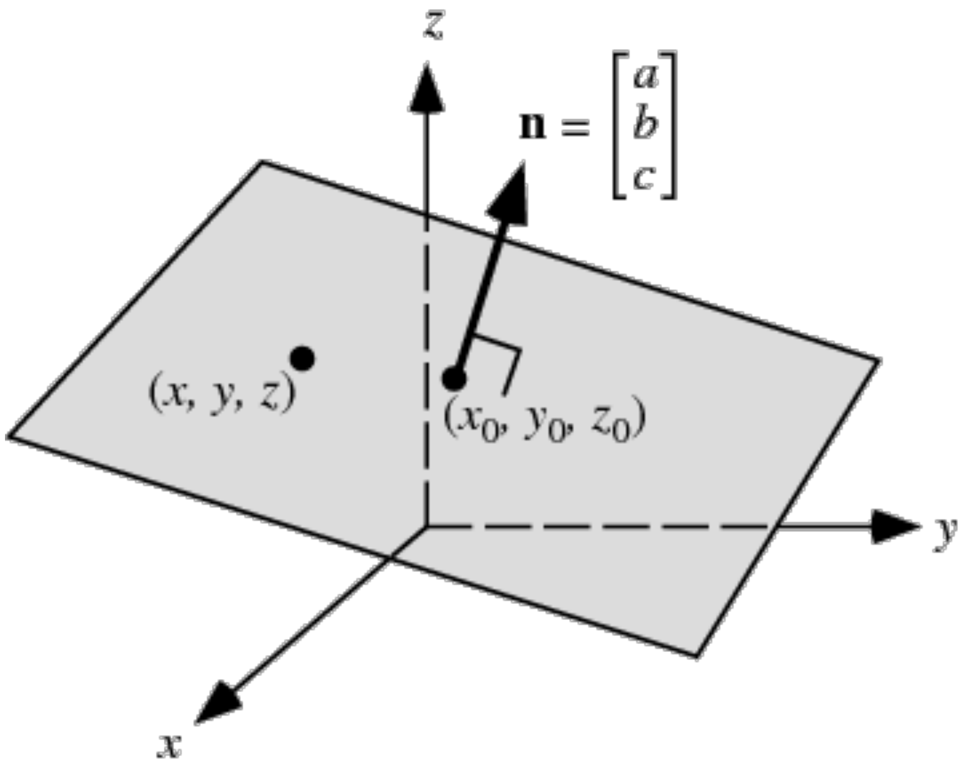
Линейный классификатор

- Добавим единичный признак

$$a(x) = \operatorname{sign} \sum_{j=1}^{d+1} w_j x^j = \operatorname{sign} \langle w, x \rangle$$

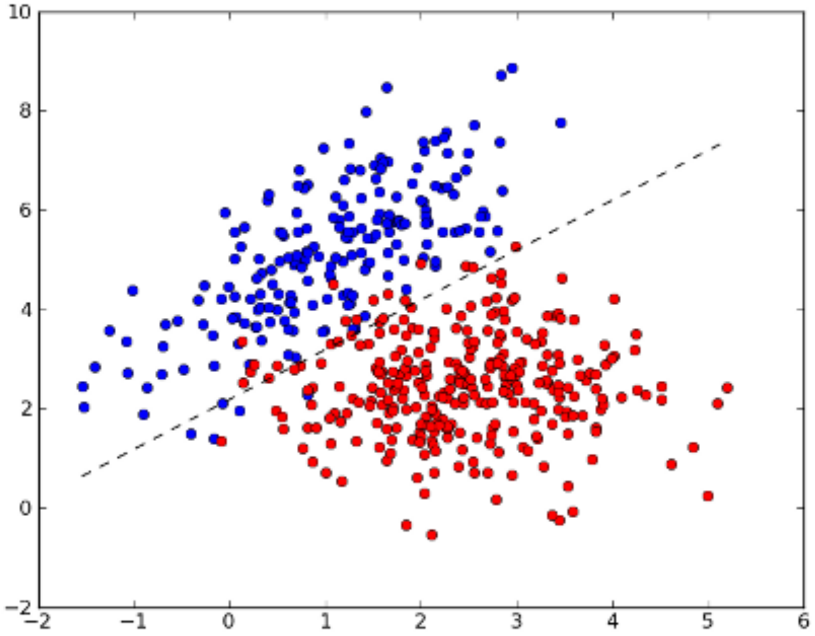
Геометрия линейного классификатора

Уравнение гиперплоскости: $\langle w, x \rangle = 0$



Геометрия линейного классификатора

- Линейный классификатор проводит гиперплоскость
- $\langle w, x \rangle < 0$ — объект «слева» от неё
- $\langle w, x \rangle > 0$ — объект «справа» от неё



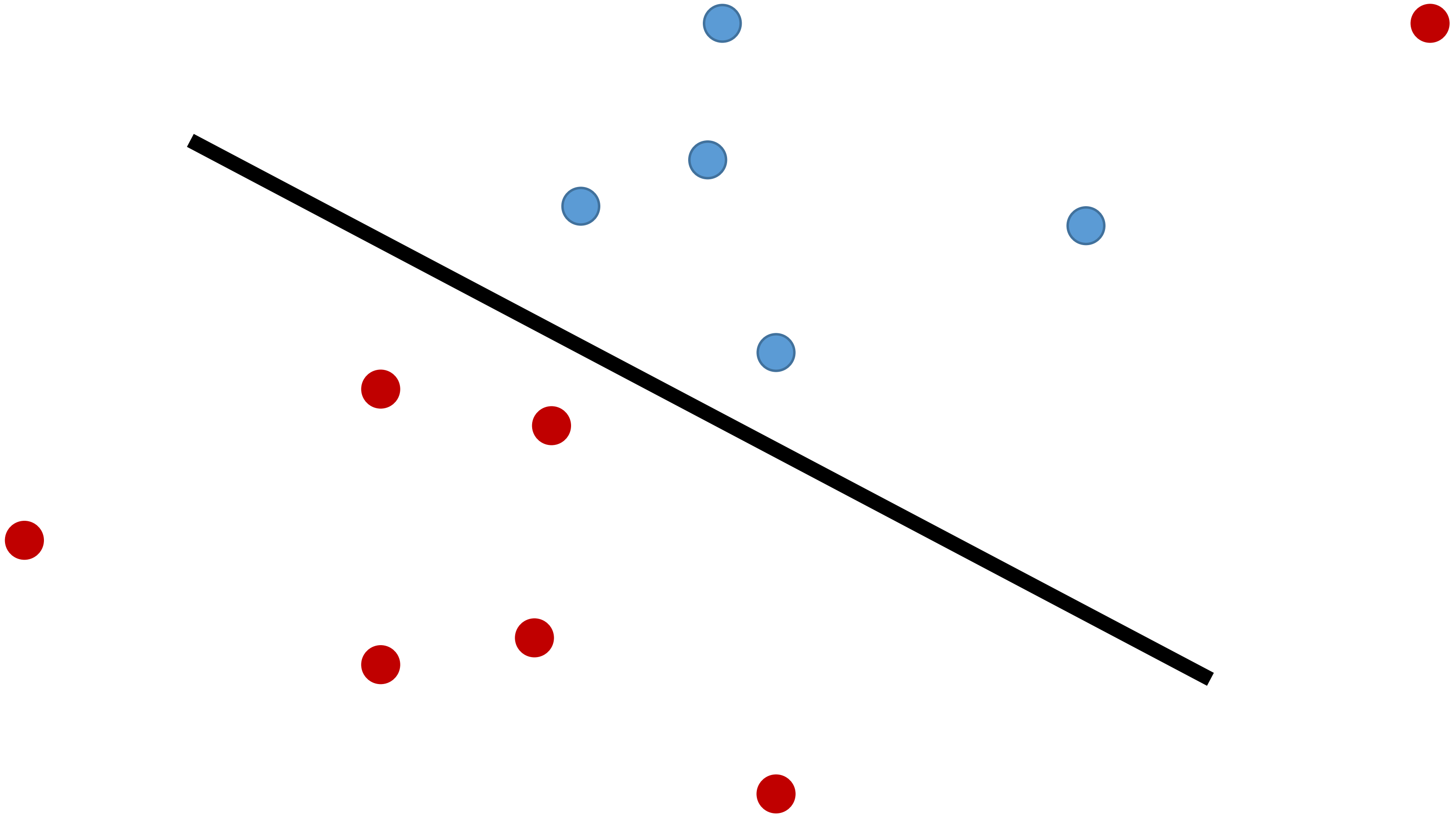
Геометрия линейного классификатора

- Расстояние от точки до гиперплоскости $\langle w, x\rangle = 0$:

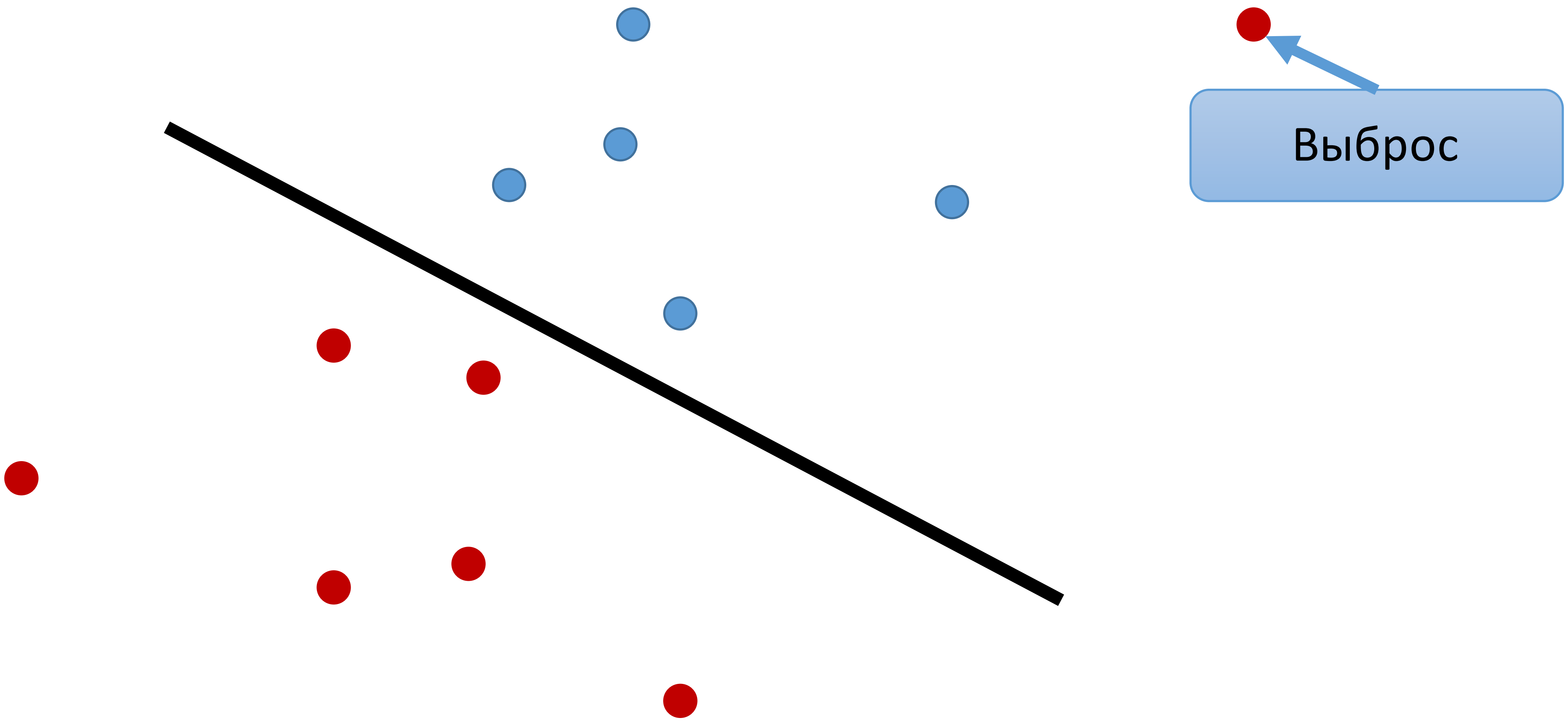
$$\frac{|\langle w, x\rangle|}{\|w\|}$$

- Чем больше $\langle w, x\rangle$, тем дальше объект от разделяющей гиперплоскости

Геометрия линейного классификатора

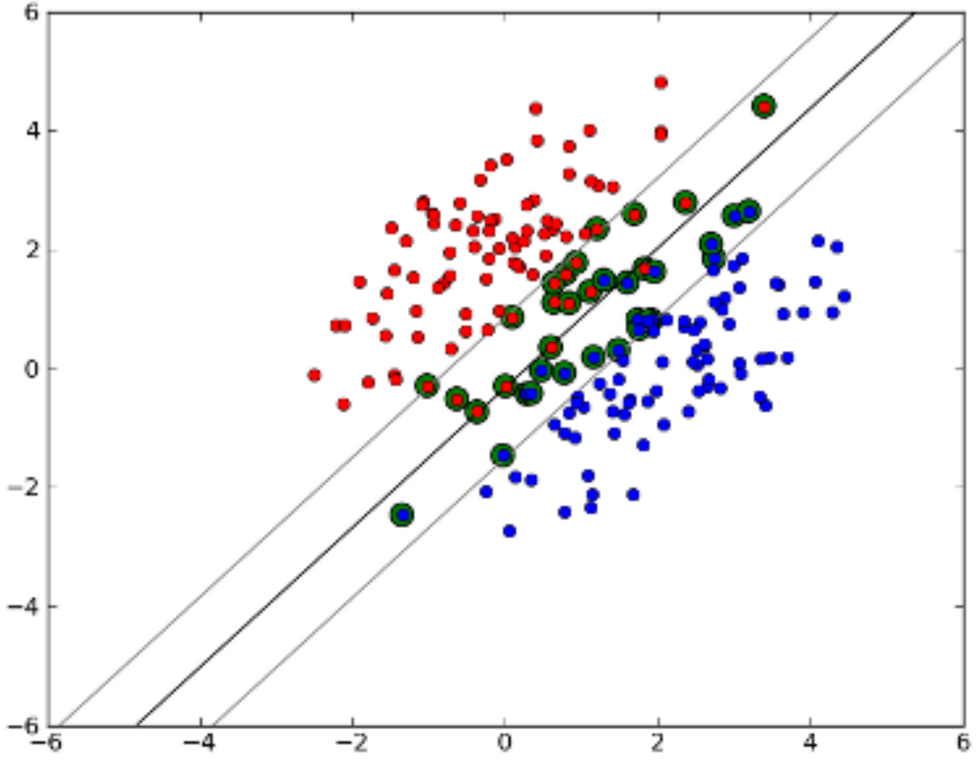


Геометрия линейного классификатора



Отступ

- $M_i = y_i \langle w, x_i \rangle$
- $M_i > 0$ — классификатор дает верный ответ
- $M_i < 0$ — классификатор ошибается
- Чем дальше отступ от нуля, тем больше уверенности



Линейный классификатор

- Линейный классификатор разделяет два класса гиперплоскостью
- Чем больше отступ по модулю, тем дальше объект от гиперплоскости
- Знак отступа говорит о корректности предсказания

Метрики качества классификации

Качество классификации

- Доля правильных ответов (accuracy):

$$\frac{1}{\ell}\sum_{i=1}^{\ell}[a(x_i) = y_i]$$

Улучшение метрики

- Два алгоритма
- Доли правильных ответов: r_1 и r_2
- Абсолютное улучшение: $r_2 - r_1$
- Относительное улучшение: $\frac{r_2 - r_1}{r_1}$

Улучшение метрики

- $r_1 = 0.8$
- $r_2 = 0.9$
- $\frac{r_2 - r_1}{r_1} = 12.5\%$

- $r_1 = 0.5$
- $r_2 = 0.75$
- $\frac{r_2 - r_1}{r_1} = 50\%$

- $r_1 = 0.001$
- $r_2 = 0.01$
- $\frac{r_2 - r_1}{r_1} = 900\%$

Матрица ошибок

	$\boldsymbol{y = 1}$	$\boldsymbol{y = -1}$
$a(x) = 1$	True Positive (TP)	False Positive (FP)
$a(x) = -1$	False Negative (FN)	True Negative (TN)

Точность (precision)

- Можно ли доверять классификатору при $a(x) = 1$?

$$\text{precision}(a, X) = \frac{TP}{TP + FP}$$

Полнота (recall)

- Как много положительных объектов находит классификатор?

$$\text{recall}(a, X) = \frac{TP}{TP + FN}$$

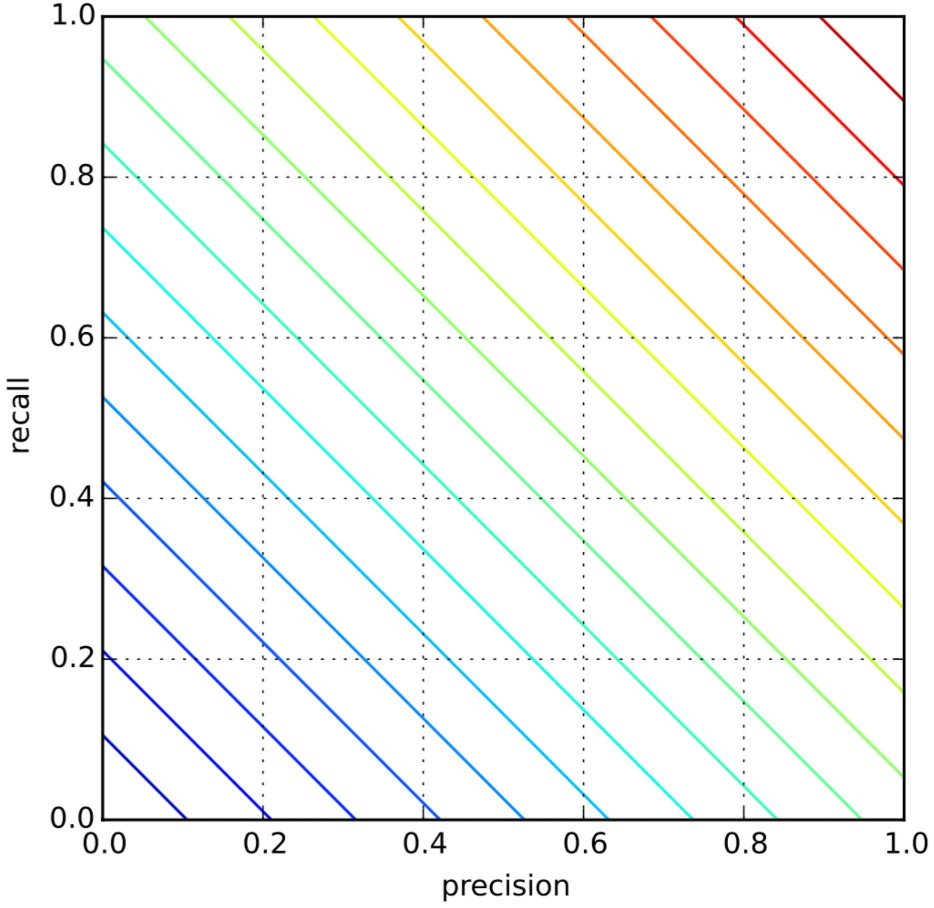
Точность и полнота

- Точность — можно ли доверять классификатору при $a(x) = 1$?
- Полнота — как много положительных объектов находит $a(x)$?

- Оптимизировать две метрики одновременно очень неудобно
- Как объединить?

Арифметическое среднее

$$A = \frac{1}{2}(\text{precision} + \text{recall})$$

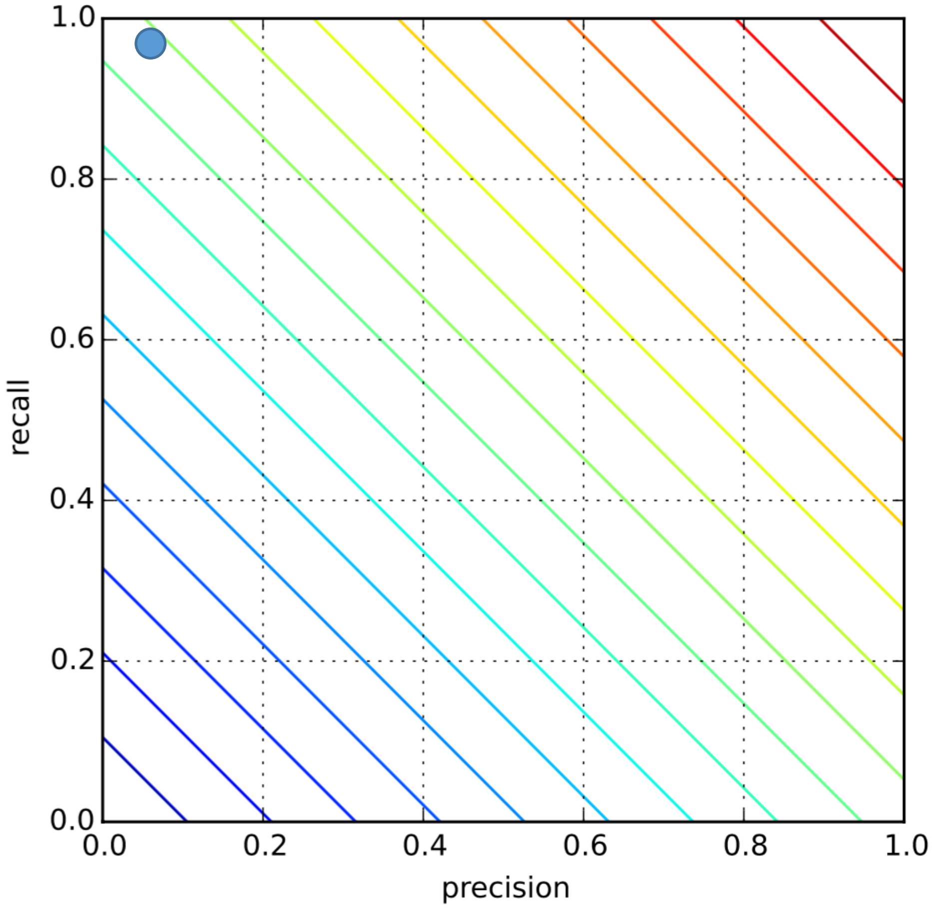


Арифметическое среднее

$$A = \frac{1}{2}(\text{precision} + \text{recall})$$

- $\text{precision} = 0.1$
- $\text{recall} = 1$
- $A = 0.55$

- Плохой алгоритм

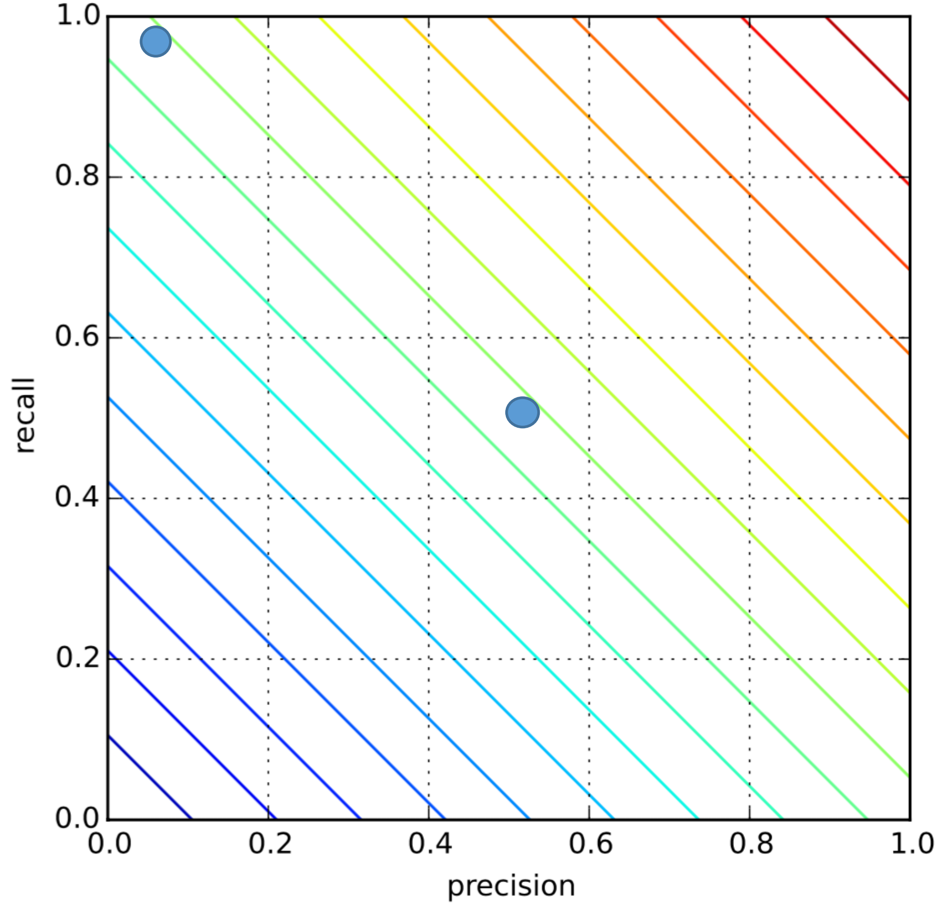


Арифметическое среднее

$$A = \frac{1}{2}(\text{precision} + \text{recall})$$

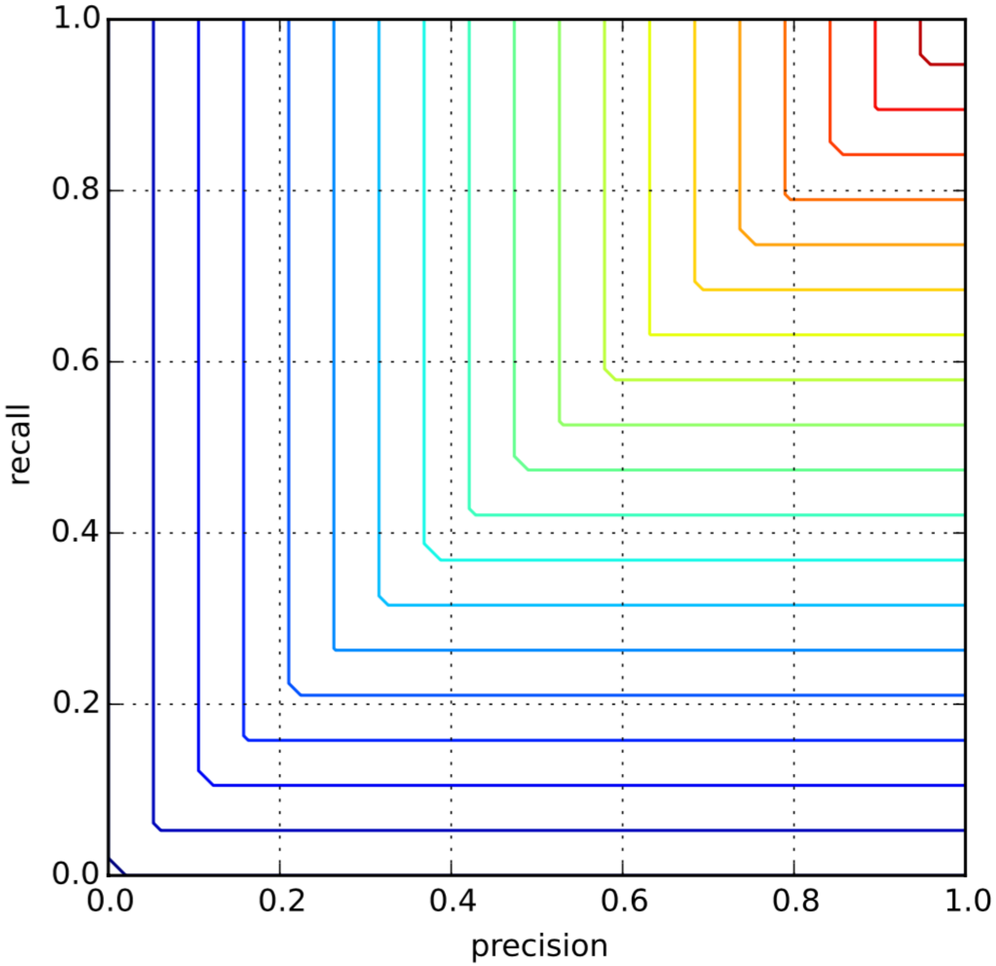
- $\text{precision} = 0.55$
- $\text{recall} = 0.55$
- $A = 0.55$

- Нормальный алгоритм
- Но качество такое же, как у плохого



Минимум

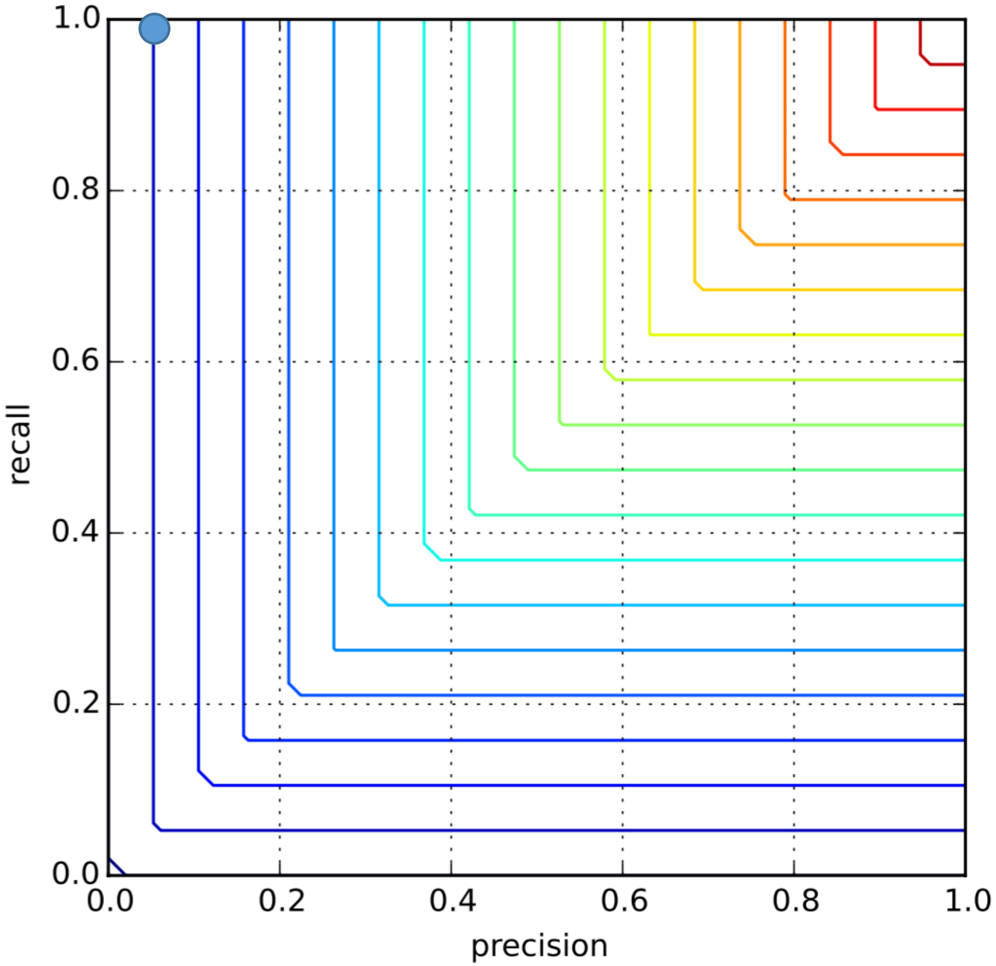
$$M = \min(\text{precision}, \text{recall})$$



Минимум

$$M = \min(\text{precision}, \text{recall})$$

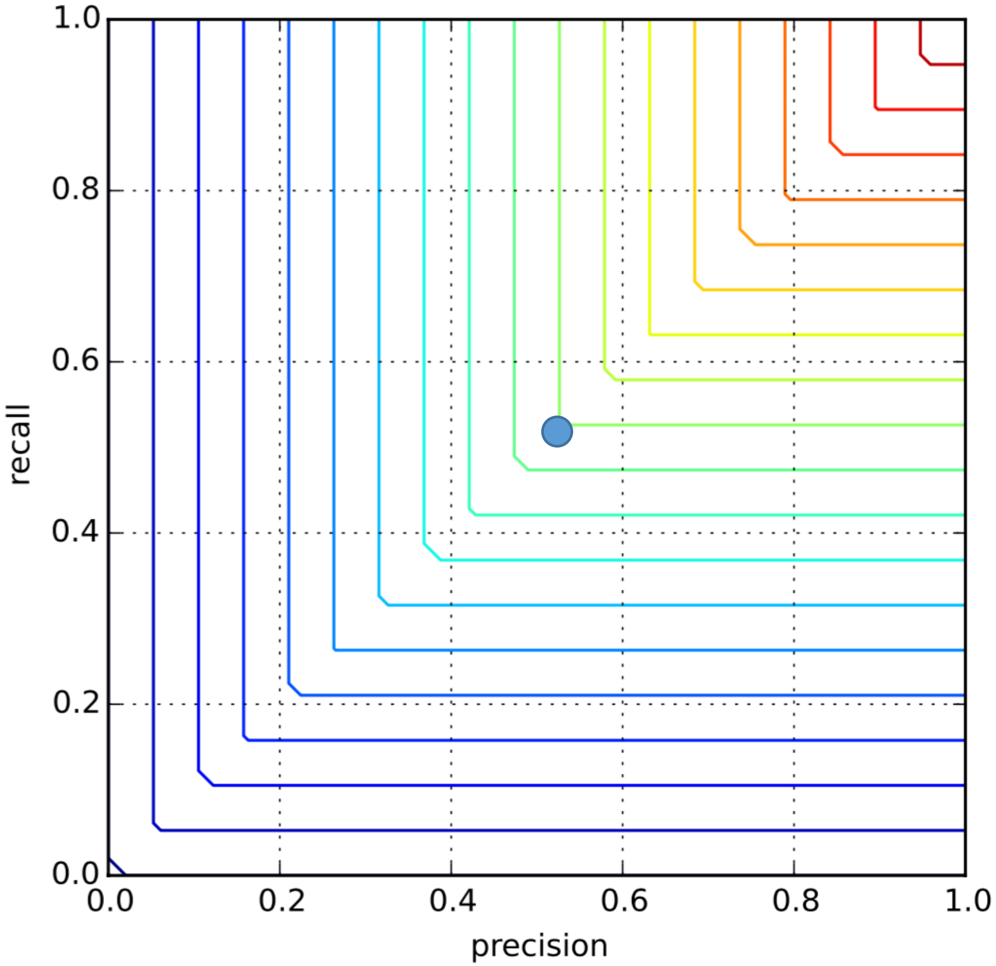
- $\text{precision} = 0.05$
- $\text{recall} = 1$
- $M = 0.05$



Минимум

$$M = \min(\text{precision}, \text{recall})$$

- $\text{precision} = 0.55$
- $\text{recall} = 0.55$
- $M = 0.55$



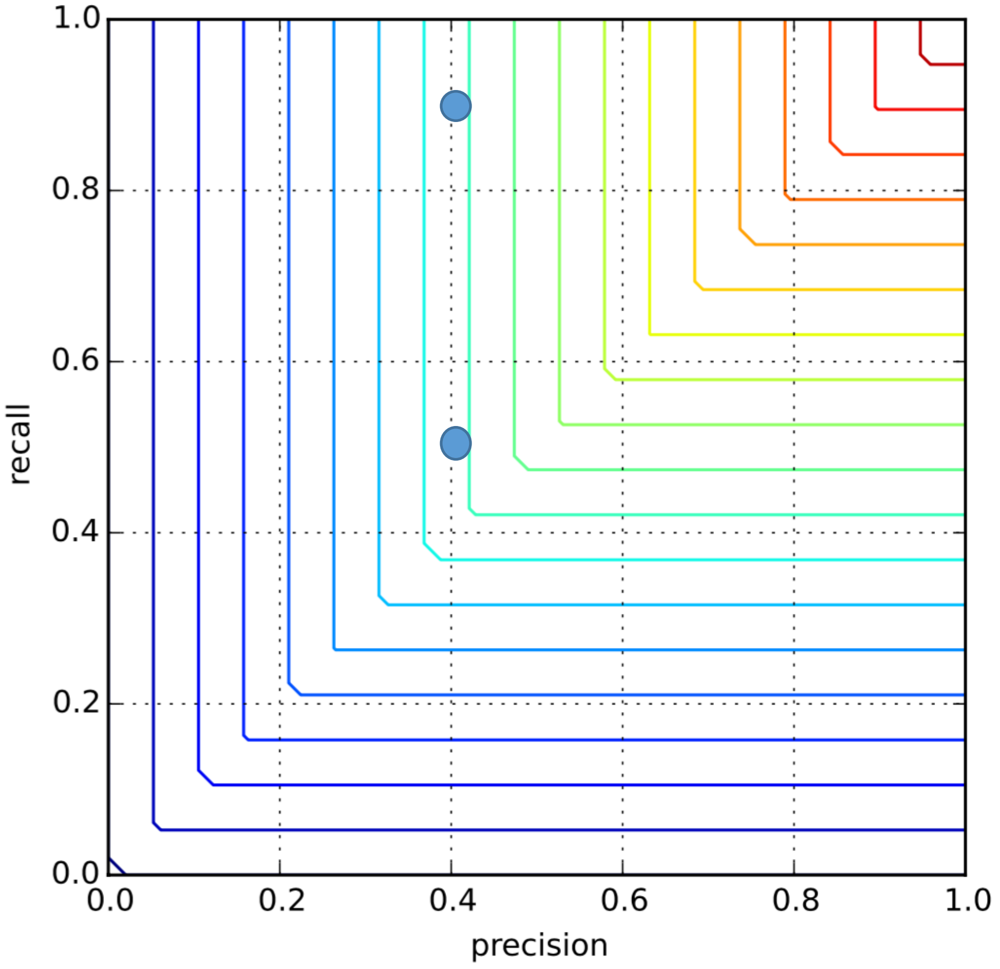
Минимум

$$M = \min(\text{precision}, \text{recall})$$

- $\text{precision} = 0.4$, $\text{recall} = 0.5$
- $M = 0.4$

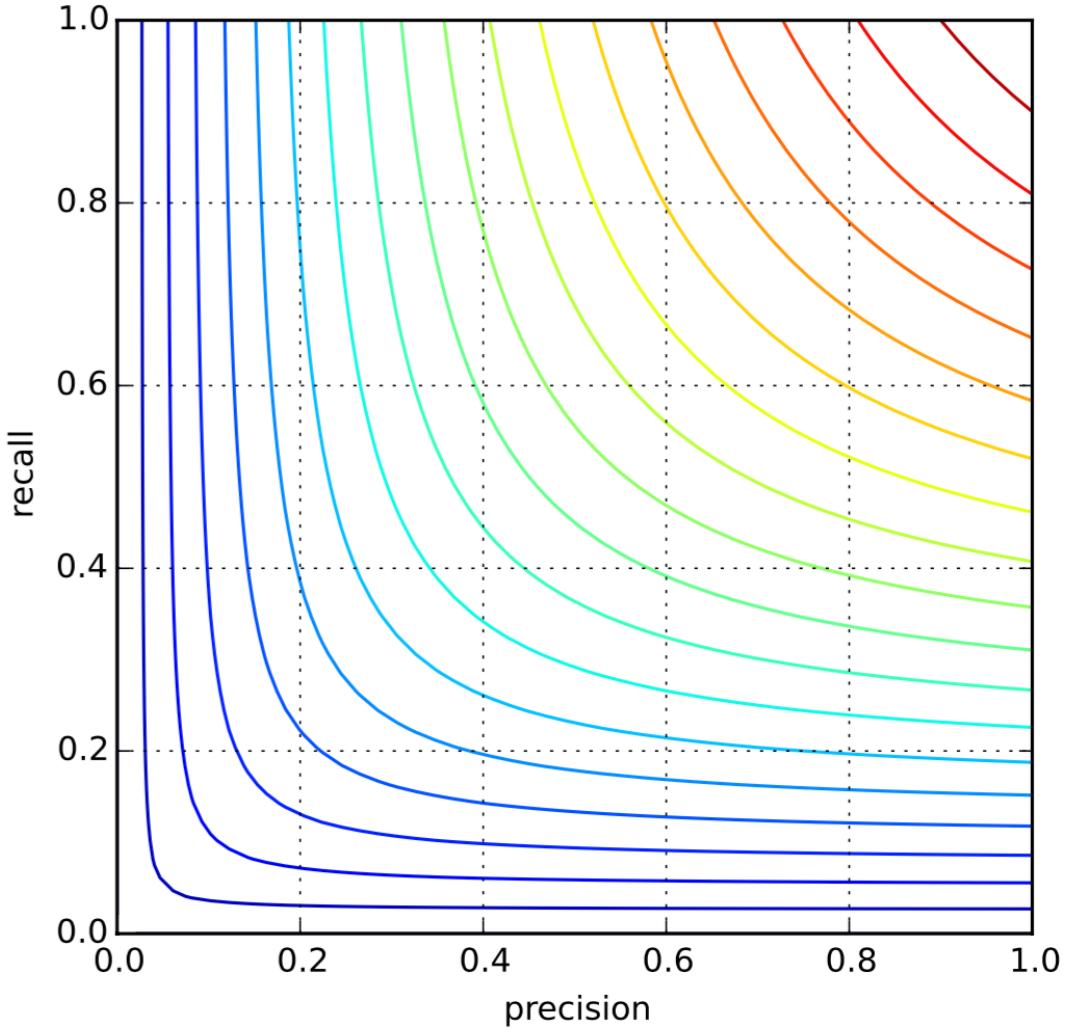
- $\text{precision} = 0.4$, $\text{recall} = 0.9$
- $M = 0.4$

- Но второй лучше!



F-мера

$$F = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

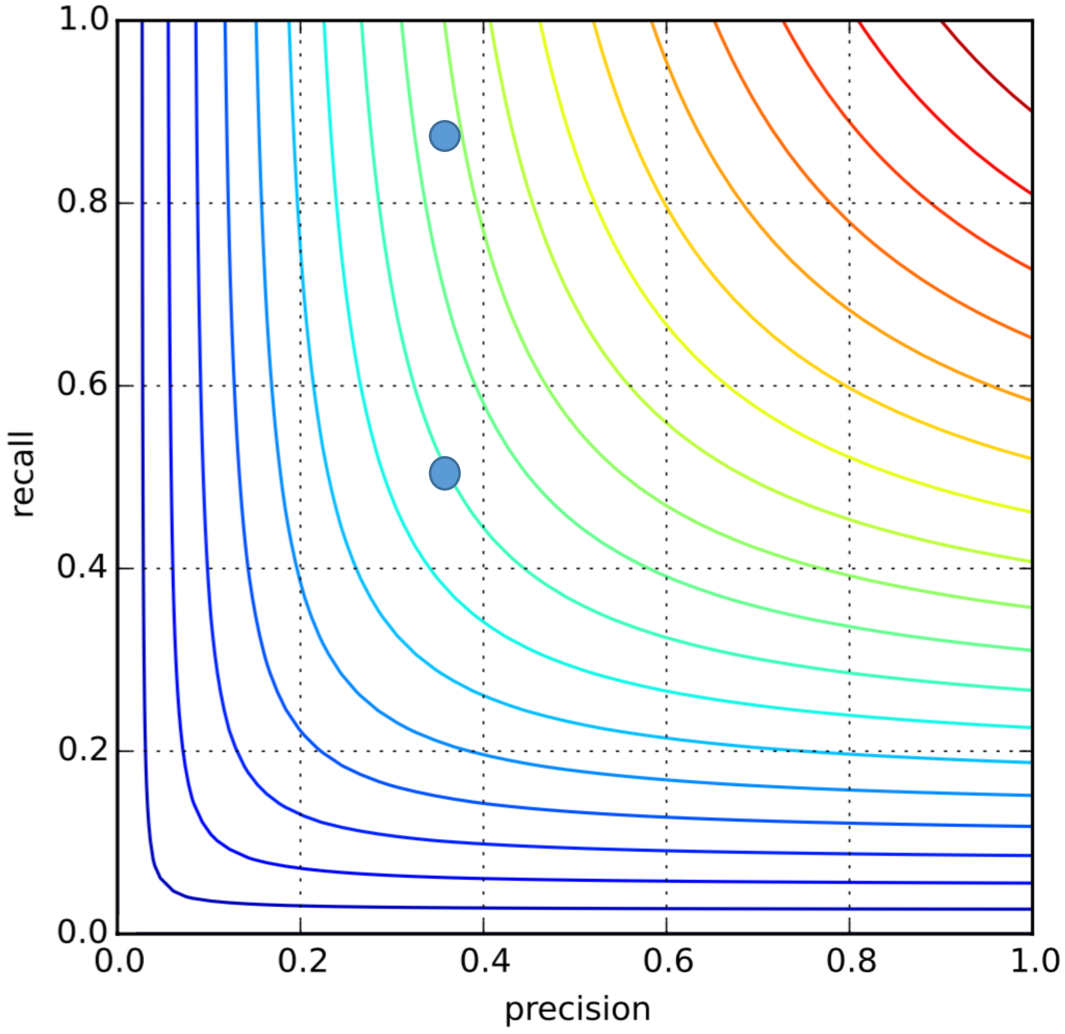


F-мера

$$F = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

- $\text{precision} = 0.4, \text{recall} = 0.5$
- $F = 0.44$

- $\text{precision} = 0.4, \text{recall} = 0.9$
- $M = 0.55$



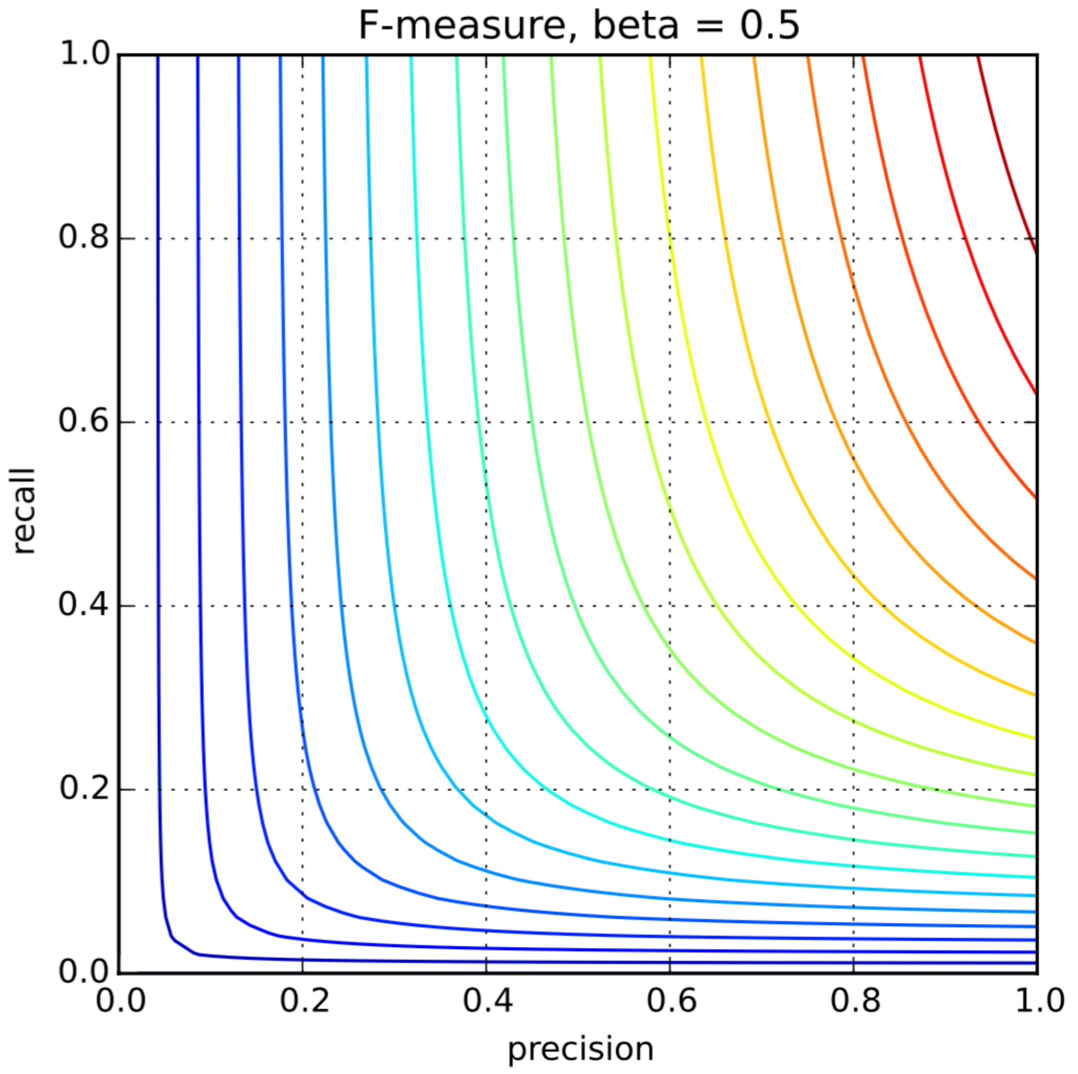
F-мера

$$F_\beta = (1+\beta^2)\frac{\text{precision} * \text{recall}}{\beta^2 * \text{precision} + \text{recall}}$$

F-мера

$$F_\beta = (1+\beta^2)\frac{\text{precision} * \text{recall}}{\beta^2 * \text{precision} + \text{recall}}$$

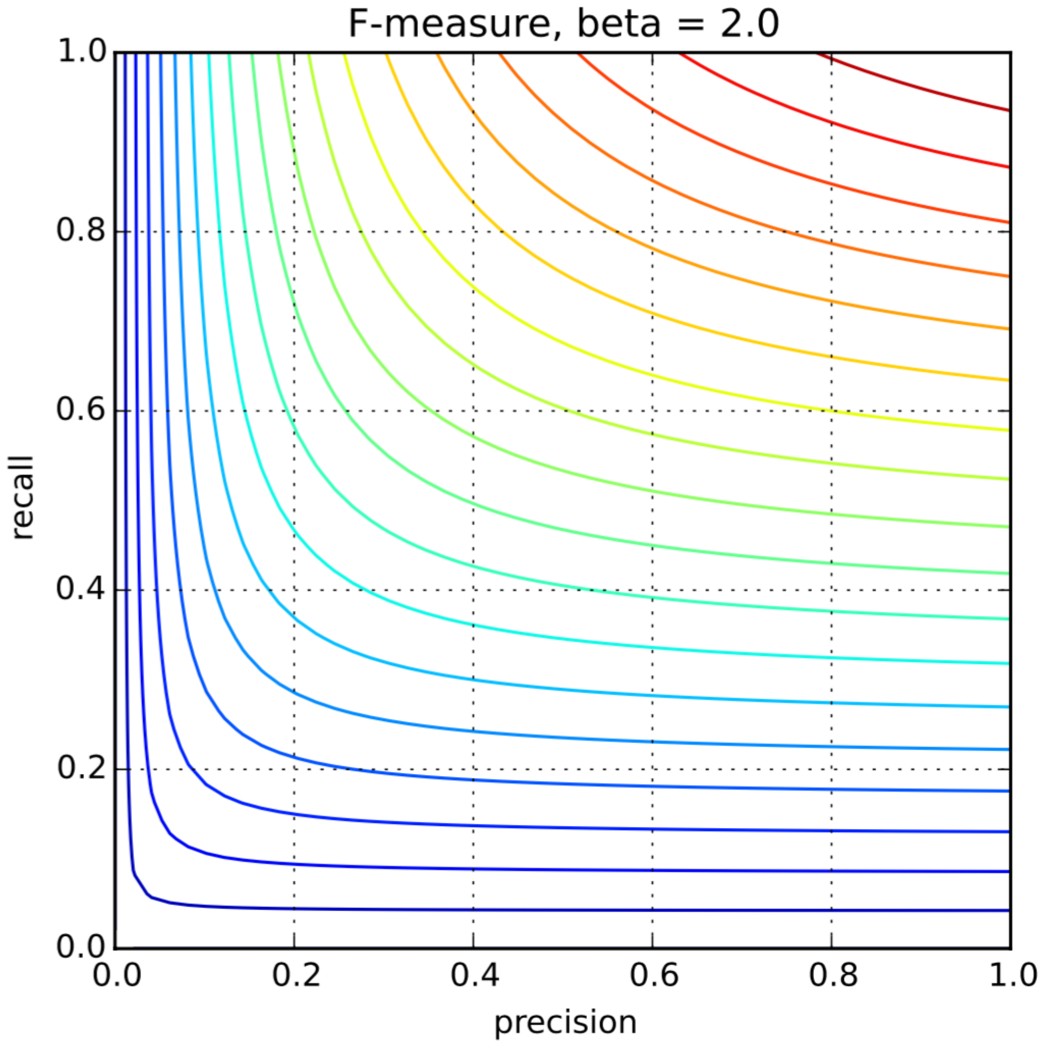
- $\beta = 0.5$
- Важнее полнота



F-мера

$$F_\beta = (1+\beta^2)\frac{\text{precision} * \text{recall}}{\beta^2 * \text{precision} + \text{recall}}$$

- $\beta = 2$
- Важнее точность



Оценки принадлежности классу

Классификатор

- Частая ситуация:

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

- $b(x)$ — оценка принадлежности классу +1

Линейный классификатор

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

- $b(x) = \langle w, x \rangle$ — оценка принадлежности классу +1
- Обычно $t = 0$

Оценка принадлежности

- Как оценить качество $b(x)$?
- Порог выбирается позже
- Порог зависит от ограничений на точность или полноту

Оценка принадлежности

- Высокий порог:
  - Мало объектов относим к +1
  - Точность выше
  - Полнота ниже
- Низкий порог:
  - Много объектов относим к +1
  - Точность ниже
  - Полнота выше

Оценка принадлежности

-1	-1	+1	-1	-1	-1	+1	+1	-1	+1
0.01	0.09	0.12	0.15	0.29	0.4	0.48	0.6	0.83	0.9

Оценка принадлежности

-1	-1	+1	-1	-1	-1	+1	+1	-1	+1
0.01	0.09	0.12	0.15	0.29	0.4	0.48	0.6	0.83	0.9

Оценка принадлежности

-1	**-1**	**+1**	**-1**	**-1**	**-1**	**+1**	**+1**	**-1**	**+1**
0.01	0.09	0.12	0.15	0.29	0.4	0.48	0.6	0.83	0.9

Оценка принадлежности

- Пример: кредитный скоринг
- $b(x)$ — оценка вероятности возврата кредита
- $a(x) = [b(x) > 0.5]$
- $\text{precision} = 0.1, \text{recall} = 0.7$
- В чем дело — в пороге или в алгоритме?

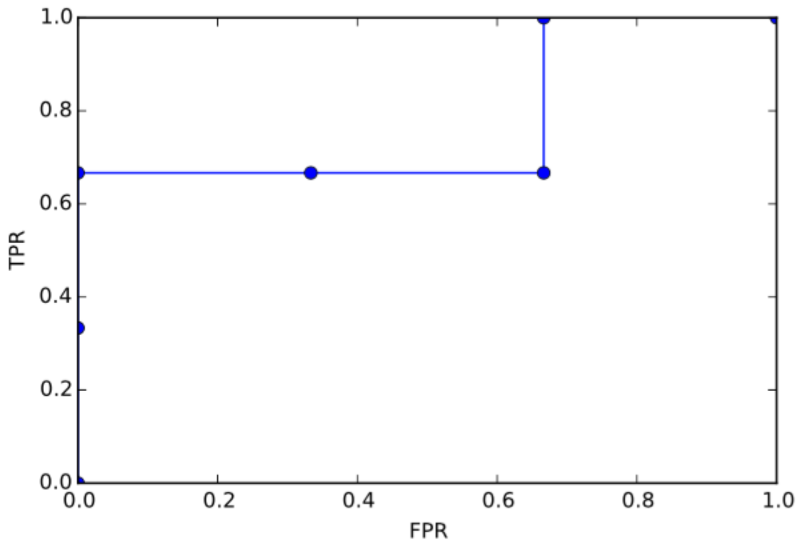
ROC-кривая

- Receiver Operating Characteristic
- Ось X — False Positive Rate

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

- Ось Y — True Positive Rate

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



ROC-кривая

- Receiver Operating Characteristic
- Ось X — False Positive Rate

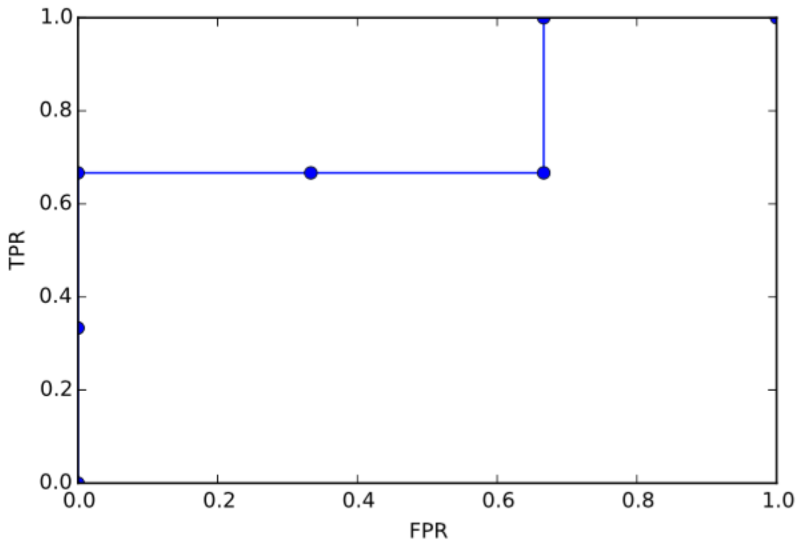
$$FPR = \frac{FP}{\boxed{FP + TN}}$$

Число отрицательных объектов

- Ось Y — True Positive Rate

$$TPR = \frac{TP}{\boxed{TP + FN}}$$

Число положительных объектов



ИНТЕГРАЛЬНАЯ МЕТРИКА: ROC-AUC

Хотим измерить качество всего семейства классификаторов независимо от выбранного порога.

Для этого будем использовать метрику AUC

AUC *– Area Under ROC Curve (площадь под ROC-кривой)*

ROC-КРИВАЯ

Для каждого значения порога t вычислим:

- False Positive Rate (доля неверно принятых объектов отрицательного класса):

$$FPR = \frac{FP}{FP + TN} = \frac{\sum_i [y_i = -1][a(x_i) = +1]}{\sum_i [y_i = -1]}$$

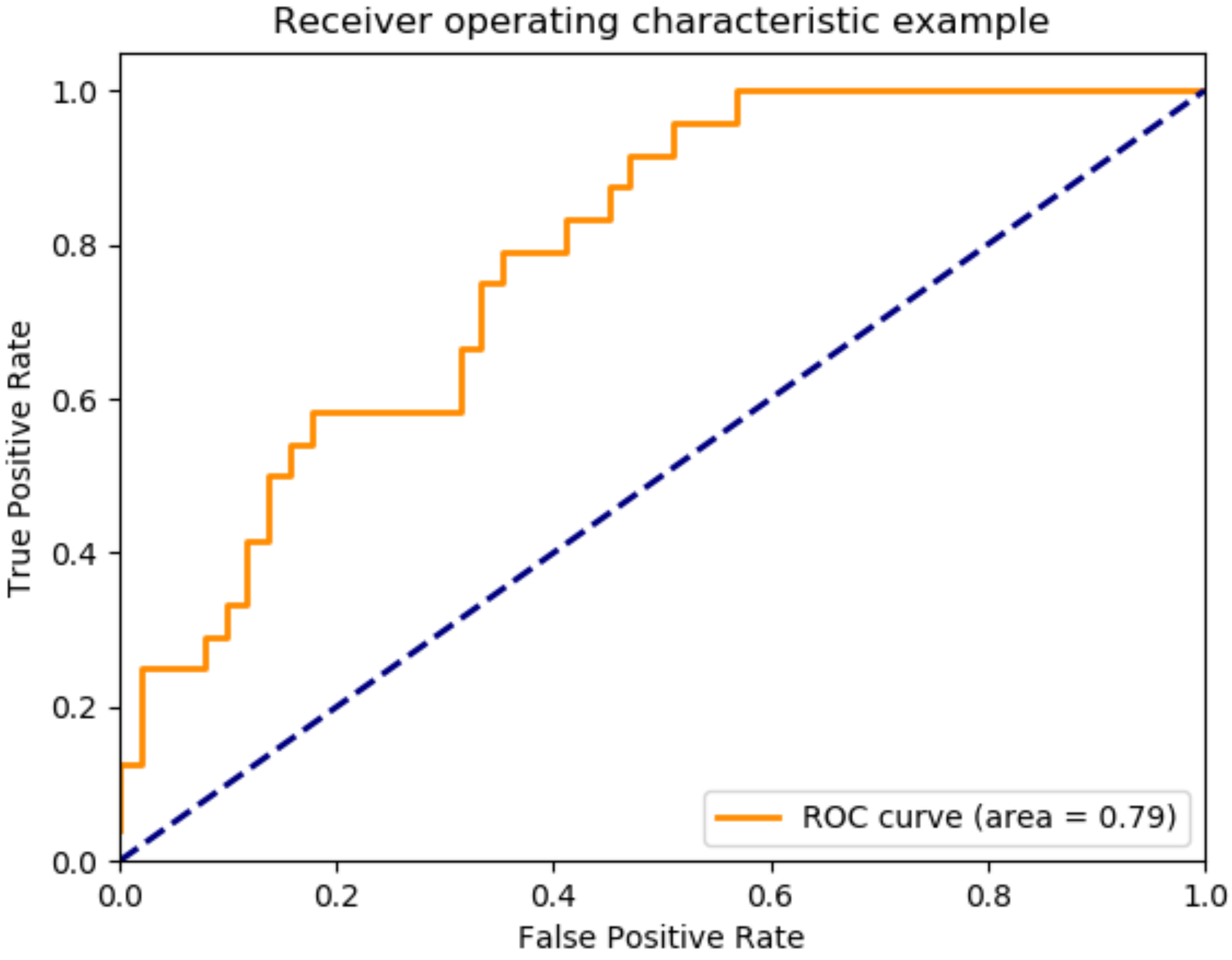
- True Positive Rate (доля верно принятых объектов положительного класса):

$$TPR = \frac{TP}{TP+FN} = \frac{\sum_i [y_i=+1]\,[a(x_i)=+1]}{\sum_i [y_i=+1]}.$$

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

ROC-КРИВАЯ

Кривая, состоящая из точек с координатами (FPR,TPR) для всех возможных порогов – это и есть ROC-кривая.

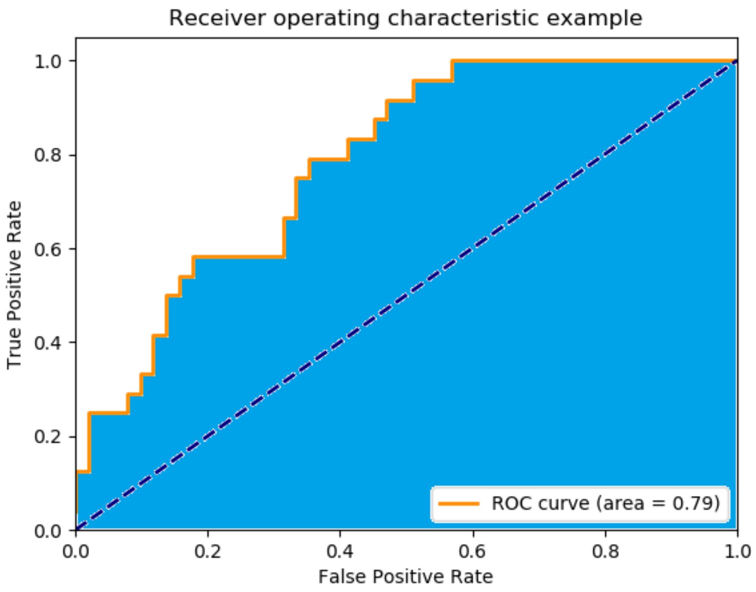


ROC-кривая. AUC.

AUC (Area Under Curve) – площадь под ROC-кривой.

$$AUC \in [0; 1].$$

- *Чему равен AUC при идеальной классификации?*
- *Чему равен AUC при случайной классификации?*



ROC-КРИВАЯ. AUC.

AUC (Area Under Curve) – площадь под ROC-кривой.

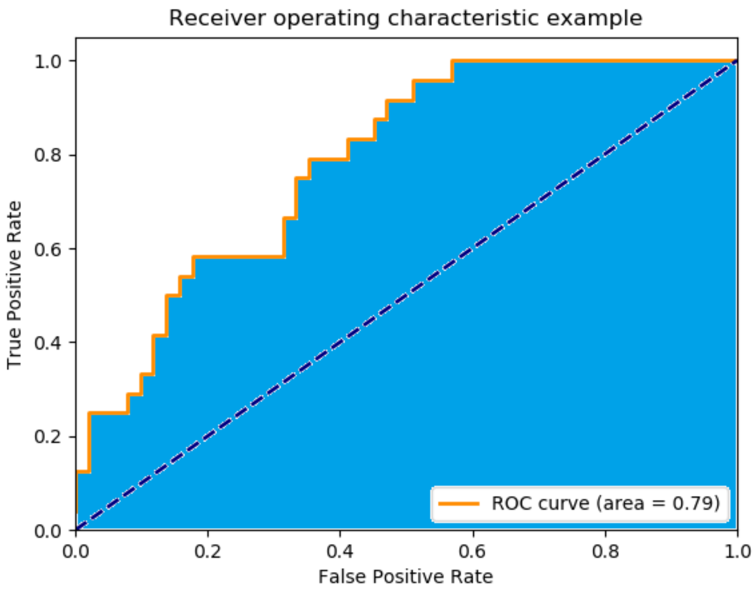
$$AUC \in [0; 1].$$

- $AUC = 1$ –

идеальная классификация

- $AUC = 0.5$ –

случайная классификация



ПРИМЕР ПОСТРОЕНИЯ ROC-КРИВОЙ

- Пусть есть выборка из 5 объектов и следующие предсказания классификатора оценки принадлежности к классу +1:

$b(x)$	0.2	0.4	0.1	0.7	0.05
y	-1	$+1$	-1	$+1$	$+1$

ПРИМЕР ПОСТРОЕНИЯ ROC-КРИВОЙ

- Пусть есть выборка из 5 объектов и следующие предсказания классификатора оценки принадлежности к классу +1:

$b(x)$	0.2	0.4	0.1	0.7	0.05
y	-1	$+1$	-1	$+1$	$+1$

- Упорядочим объекты по убыванию предсказаний:

$(0.7, 0.4, 0.2, 0.1, 0.05)$

1 шаг: $\boldsymbol{t = 0.7}$, то есть

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

$$a(x) = [b(x) > 0.7]$$

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

ПРИМЕР ПОСТРОЕНИЯ ROC-КРИВОЙ

- Пусть есть выборка из 5 объектов и следующие предсказания классификатора оценки принадлежности к классу +1:

$b(x)$	0.2	0.4	0.1	0.7	0.05
y	-1	$+1$	-1	$+1$	$+1$

- Упорядочим объекты по убыванию предсказаний:

$(0.7, 0.4, 0.2, 0.1, 0.05)$

1 шаг: $\boldsymbol{t = 0.7}$, то есть

$$\boldsymbol{TPR = \frac{TP}{TP+FN}}$$

$$a(x) = [b(x) > 0.7]$$

$$\boldsymbol{FPR = \frac{FP}{FP+TN}}$$

$$TPR = \frac{0}{0+3} = 0,\ \ FPR = \frac{0}{0+2} = 0.$$

ПРИМЕР ПОСТРОЕНИЯ ROC-КРИВОЙ

- Пусть есть выборка из 5 объектов и следующие предсказания классификатора оценки принадлежности к классу +1:

$b(x)$	0.2	0.4	0.1	0.7	0.05
y	-1	$+1$	-1	$+1$	$+1$

- Упорядочим объекты по убыванию предсказаний:

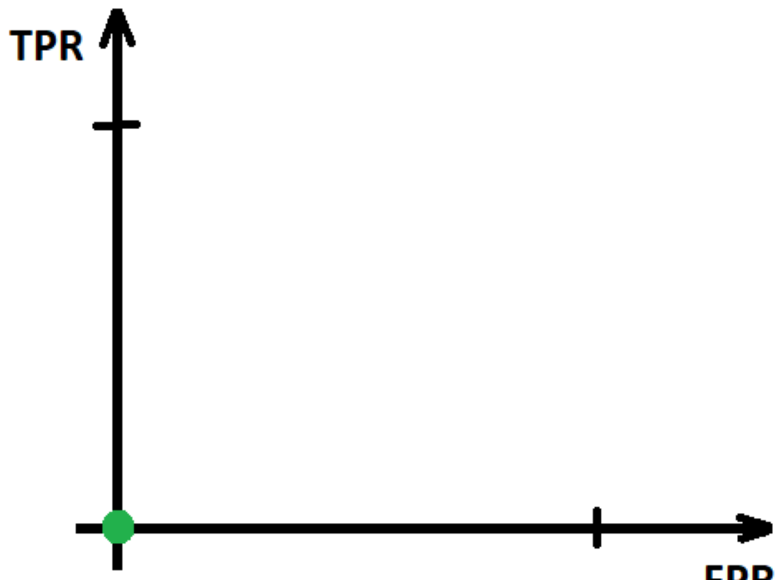
$(0.7, 0.4, 0.2, 0.1, 0.05)$

1 шаг: $\boldsymbol{t = 0.7}$, то есть

$a(x) = [b(x) > 0.7]$

$$TPR = \frac{0}{0+3} = 0,$$

$$FPR = \frac{0}{0+2} = 0.$$



ПРИМЕР ПОСТРОЕНИЯ ROC-КРИВОЙ

- Оценки принадлежности к классу +1:

$b(x)$	0.2	0.4	0.1	$\mathbf{0.7}$	0.05
y	-1	$+1$	-1	$\mathbf{+1}$	$+1$

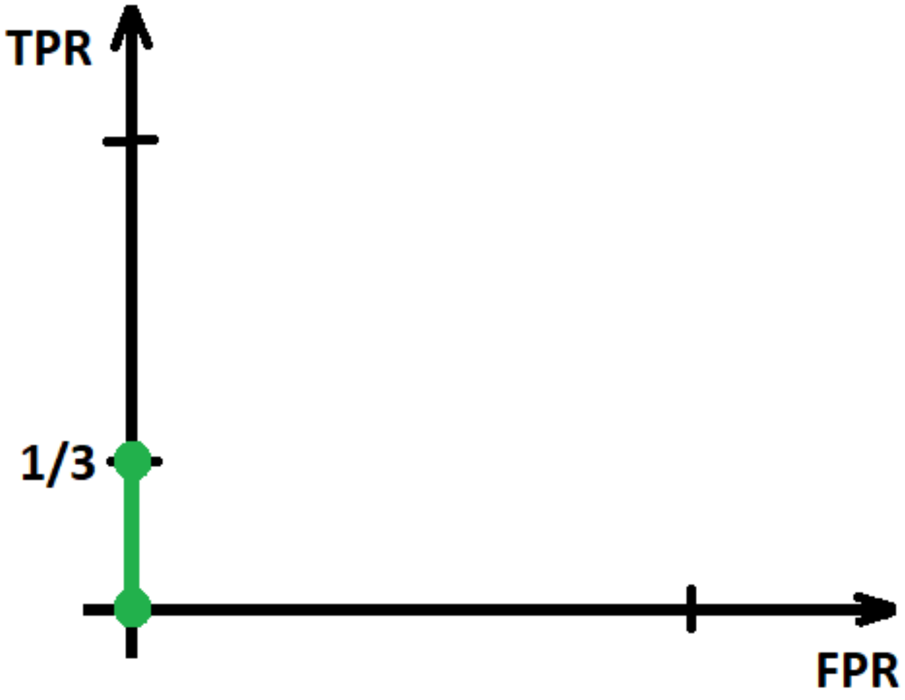
- Упорядочим объекты по

убыванию предсказаний:
$(0.7, 0.4, 0.2, 0.1, 0.05)$

2 шаг: $\boldsymbol{t = 0.4}$, то есть
$a(x) = [b(x) > 0.4]$

$TPR = \frac{1}{1+2} = \frac{1}{3}$,

$FPR = \frac{0}{0+2} = 0.$



ПРИМЕР ПОСТРОЕНИЯ ROC-КРИВОЙ

- Оценки принадлежности к классу +1:

$b(x)$	0.2	$\mathbf{0.4}$	0.1	$\mathbf{0.7}$	0.05
y	-1	$\mathbf{+1}$	-1	$\mathbf{+1}$	$+1$

- Упорядочим объекты по

убыванию предсказаний:

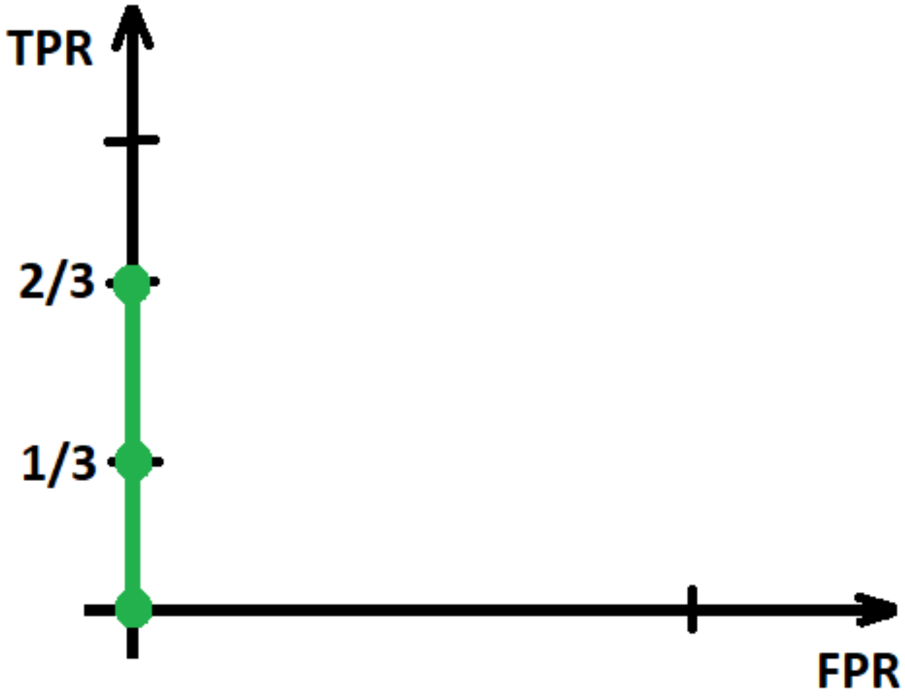
$(0.7, 0.4, 0.2, 0.1, 0.05)$

3 шаг: $\boldsymbol{t = 0.2}$, то есть

$a(x) = [b(x) > 0.2]$

$TPR = \frac{2}{2+1} = \frac{2}{3}$,

$FPR = \frac{0}{0+2} = 0.$



ПРИМЕР ПОСТРОЕНИЯ ROC-КРИВОЙ

- Оценки принадлежности к классу +1:

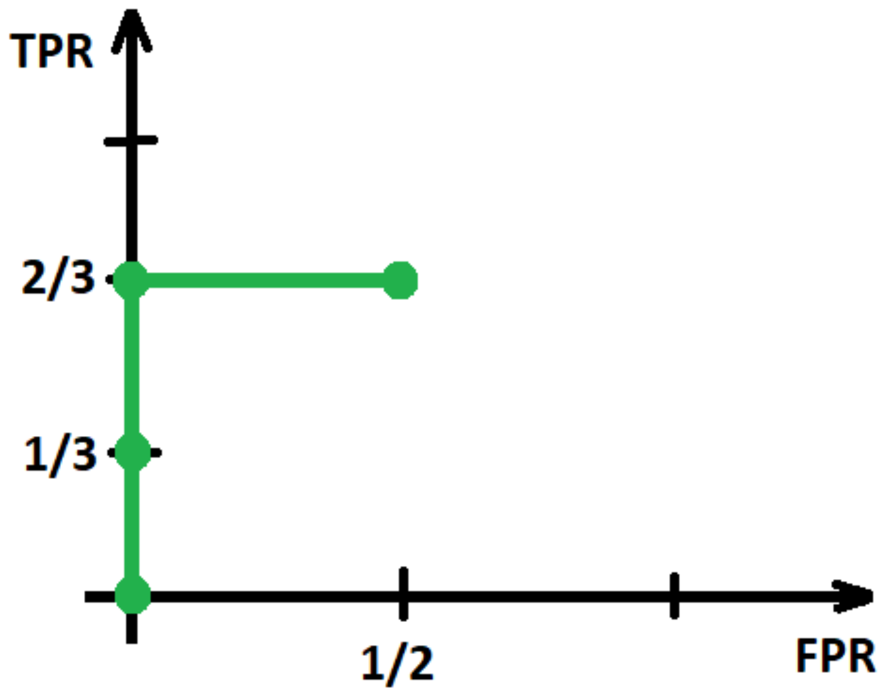
$b(x)$	$\mathbf{0.2}$	$\mathbf{0.4}$	0.1	$\mathbf{0.7}$	0.05
y	$\mathbf{-1}$	$\mathbf{+1}$	-1	$\mathbf{+1}$	$+1$

- Упорядочим объекты по убыванию предсказаний: $(0.7, 0.4, 0.2, 0.1, 0.05)$

4 шаг: $\boldsymbol{t = 0.1}$, то есть $a(x) = [b(x) > 0.1]$

$$TPR = \frac{2}{2+1} = \frac{2}{3},$$

$$FPR = \frac{1}{1+1} = \frac{1}{2}.$$



ПРИМЕР ПОСТРОЕНИЯ ROC-КРИВОЙ

- Оценки принадлежности к классу +1:

$b(x)$	$\mathbf{0.2}$	$\mathbf{0.4}$	$\mathbf{0.1}$	$\mathbf{0.7}$	0.05
y	$\mathbf{-1}$	$\mathbf{+1}$	$\mathbf{-1}$	$\mathbf{+1}$	$+1$

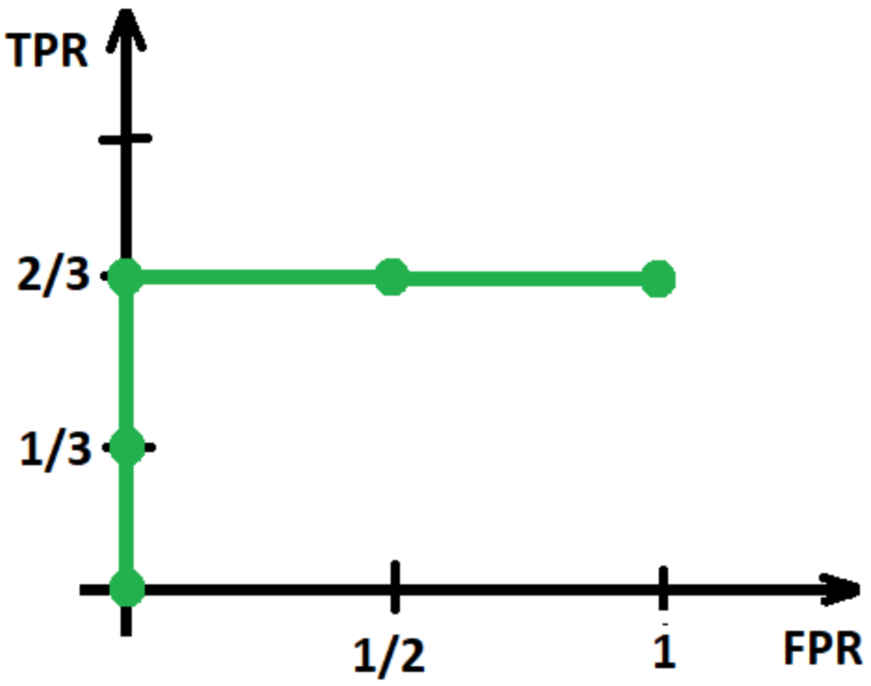
- Упорядочим объекты по

убыванию предсказаний:
$(0.7, 0.4, 0.2, 0.1, 0.05)$

5 шаг: $\boldsymbol{t = 0.05}$, то есть
$a(x) = [b(x) > 0.05]$

$TPR = \frac{2}{2+1} = \frac{2}{3}$,

$FPR = \frac{2}{2+0} = 1.$



ПРИМЕР ПОСТРОЕНИЯ ROC-КРИВОЙ

- Оценки принадлежности к классу +1:

$b(x)$	0.2	0.4	0.1	0.7	0.05
y	-1	$+1$	-1	$+1$	$+1$

- Упорядочим объекты по

убыванию предсказаний:

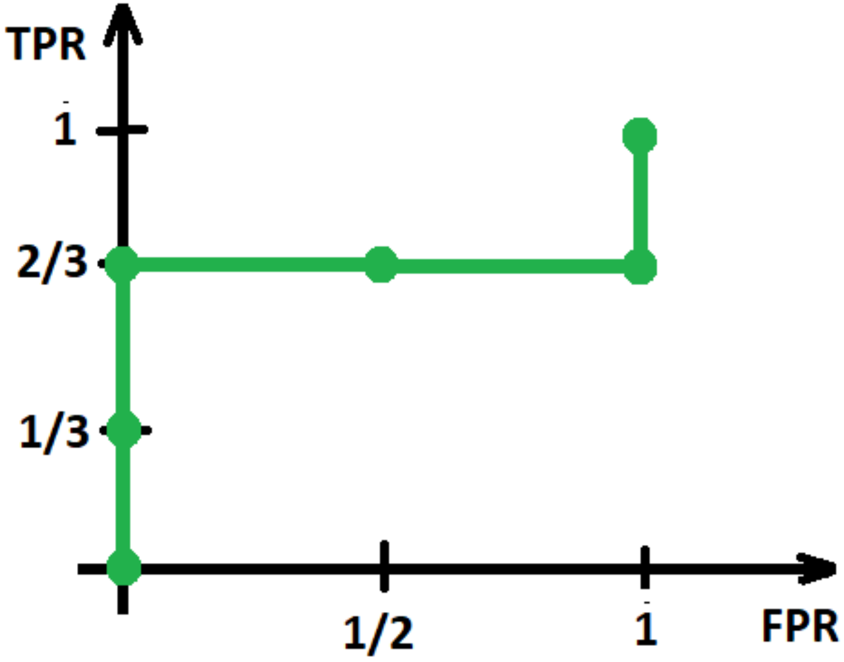
$(0.7, 0.4, 0.2, 0.1, 0.05)$

6 шаг: $\boldsymbol{t = 0}$, то есть

$a(x) = [b(x) > 0]$

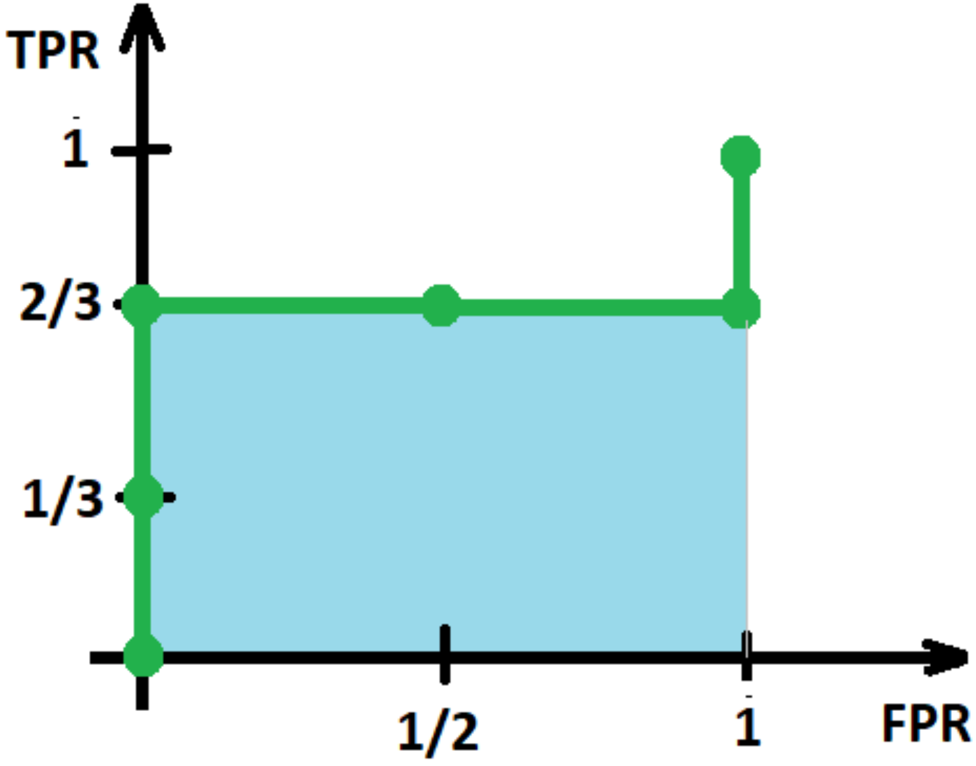
$TPR = \frac{3}{3+0} = 1,$

$FPR = \frac{2}{2+0} = 1.$



ПРИМЕР ПОСТРОЕНИЯ ROC-КРИВОЙ

$$AUC = 2/3$$

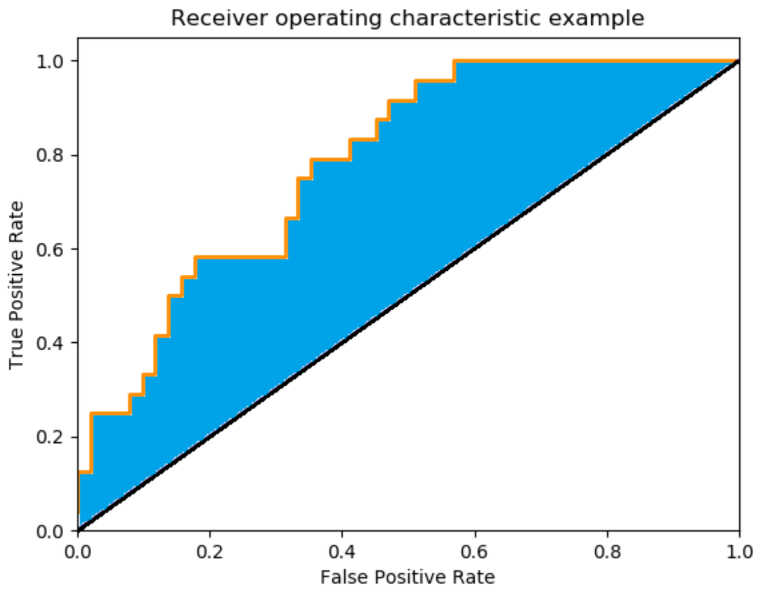


ИНДЕКС ДЖИНИ

Индекс Джини:

$$Gini = 2 \cdot AUC - 1$$

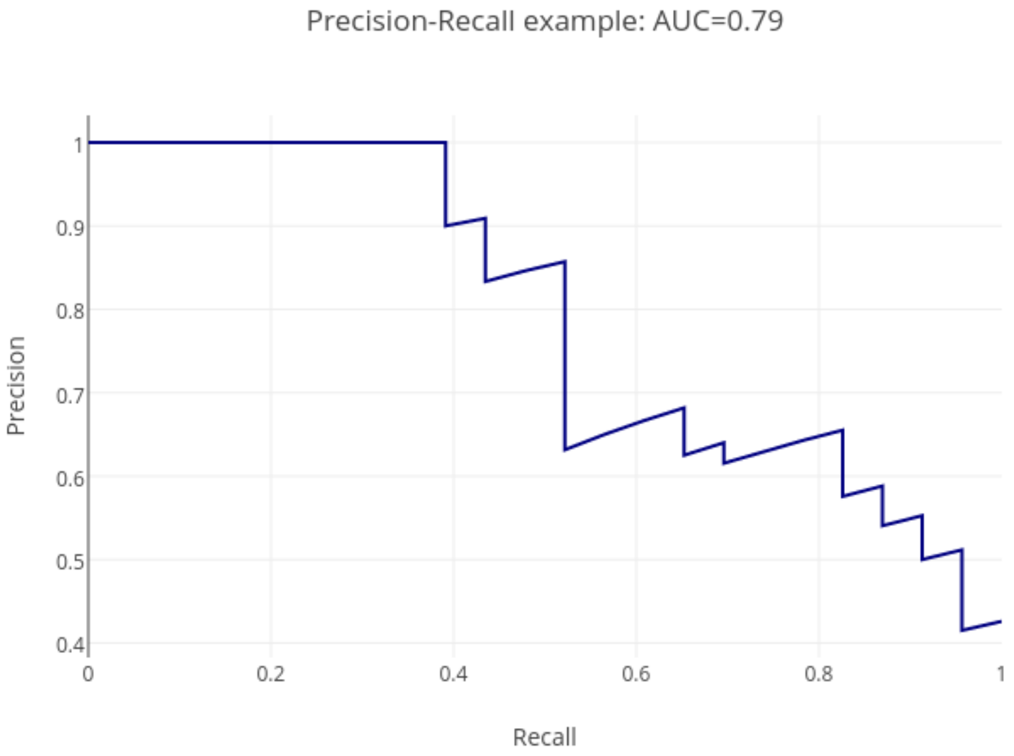
- Индекс Джини – это удвоенная площадь между главной диагональю и ROC-кривой.



PRECISION-RECALL КРИВАЯ

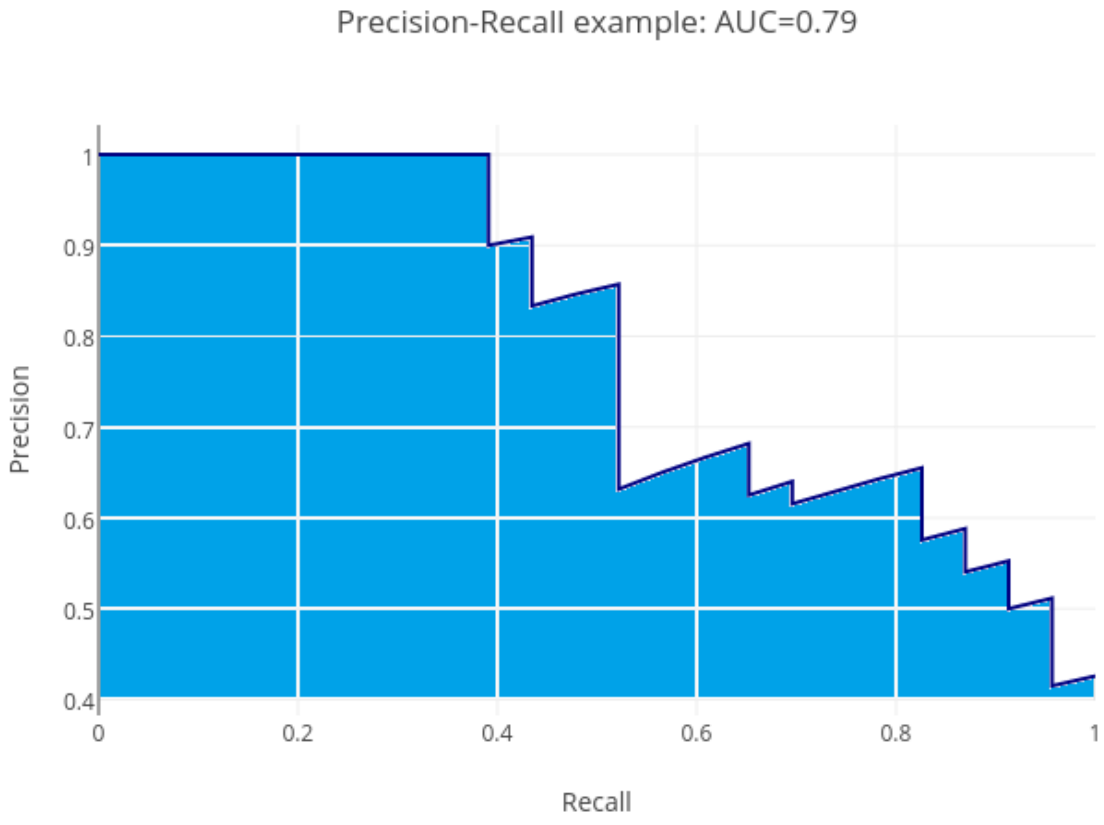
- В случае малой доли объектов положительного класса AUC-ROC может давать неадекватно хороший результат

Precision-Recall кривая:



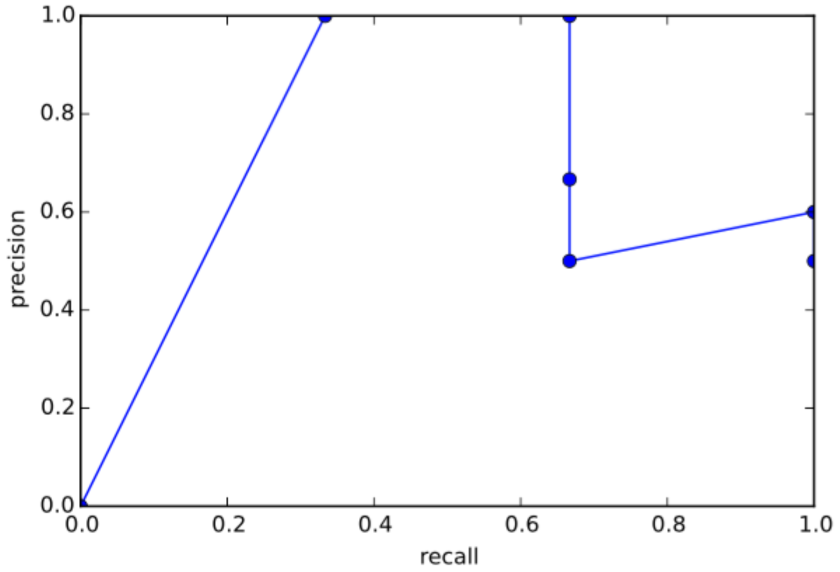
AUC-PR

AUC-PR – площадь под PR-кривой

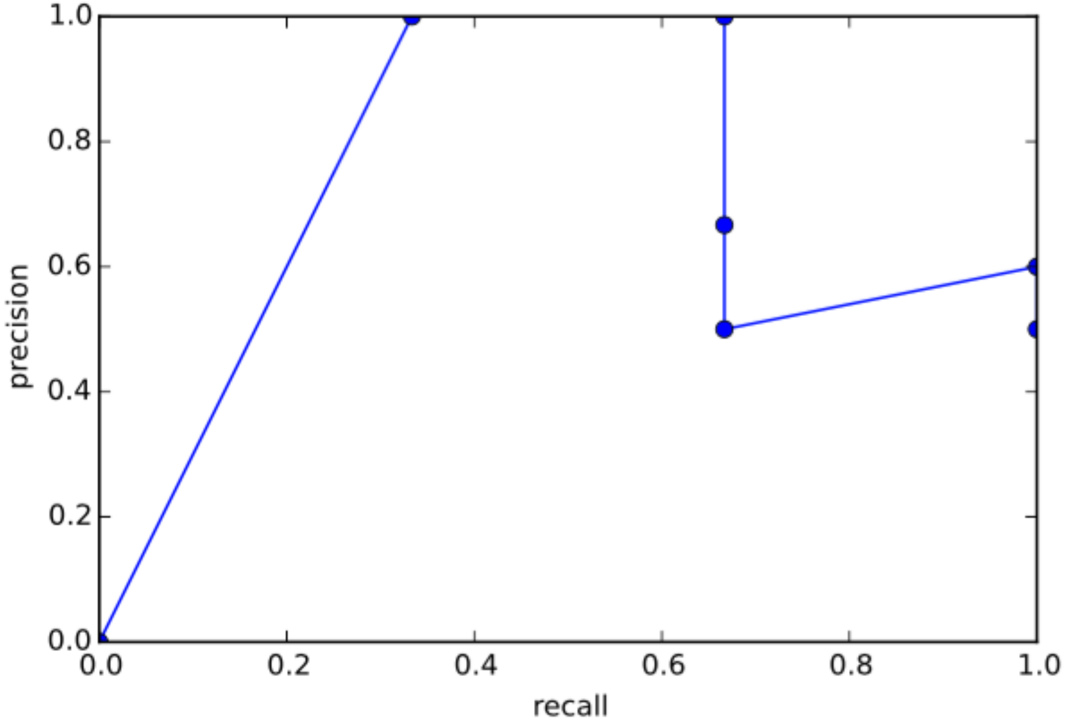


PR-кривая

- Кривая точности-полноты
- Ось X — полнота
- Ось Y — точность
- Точки — значения точности и полноты при последовательных порогах



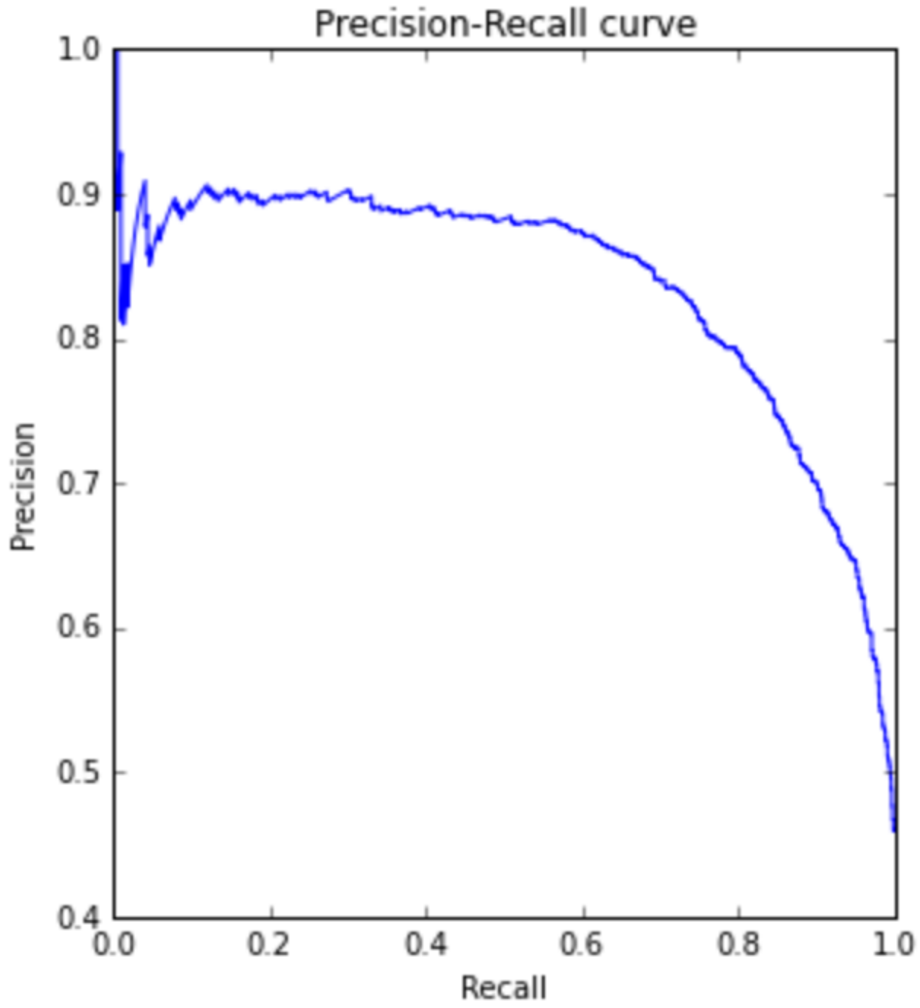
PR-кривая



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

PR-кривая

- Левая точка: (0, 0)
- Правая точка: $(1, r)$, r — доля положительных объектов
- Для идеального классификатора проходит через (1, 1)
- AUC-PRC — площадь под PR-кривой



PR-кривая в реальности