СБЕР УНИВЕРСИТЕТ

ML. Второе занятие

План занятия

1. Линейная классификация
2. Обучение линейного классификатора
3. Измерение ошибки в задачах классификации. Метрики качества
4. Метрики качества ранжирования
5. Логистическая регрессия
6. Метод опорных векторов
7. Многоклассовая классификация
8. Метрики качества многоклассовой классификации
9. Классификация с пересекающимися классами

Линейная классификация

$y_1, y_2, \ldots, y_l \in \{1, 2, \ldots, k\}$

Линейная классификация

$y_1, y_2, \ldots, y_l \in \{1, 2, \ldots, k\}$

Бинарная классификация
$y_1, y_2, \ldots, y_l \in \{-1, 1\}$

Линейная классификация

$y_1, y_2, \dots, y_l \in \{1, 2, \dots, k\}$

Бинарная классификация
$y_1, y_2, \dots, y_l \in \{-1, 1\}$

$a(x) = sign(< w, x > \ + w_0)$

Линейная классификация

$y_1, y_2, \dots, y_l \in \{1, 2, \dots, k\}$

Бинарная классификация
$y_1, y_2, \dots, y_l \in \{-1, 1\}$

$a(x) = sign(< w, x > \ + w_0)$

$< w, x > \ + w_0 = 0$ - гиперплоскость

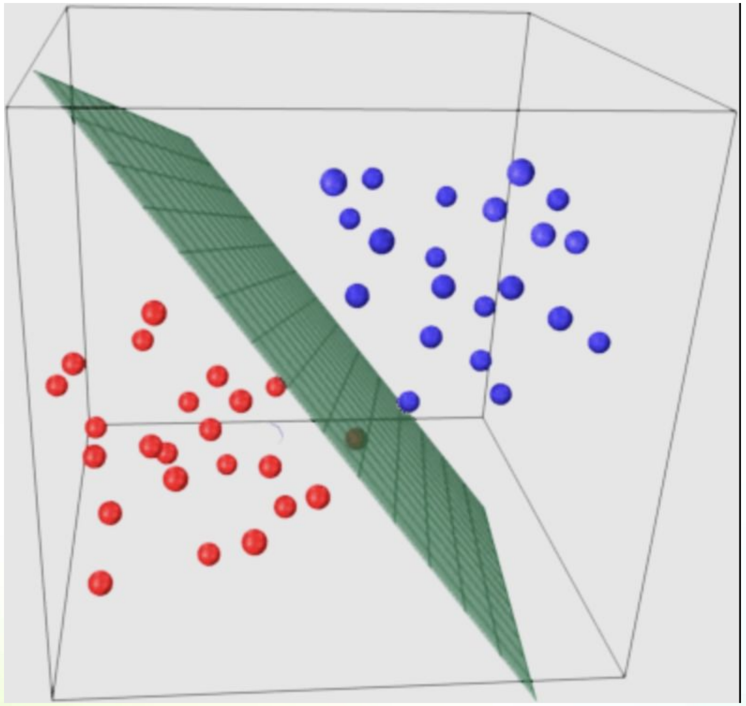
Линейная классификация

$y_1, y_2, \dots, y_l \in \{1, 2, \dots, k\}$

Бинарная классификация
$y_1, y_2, \dots, y_l \in \{-1, 1\}$

$a(x) = sign(< w, x > \ + w_0)$

$< w, x > \ + w_0 = 0$ - гиперплоскость



Обучение линейного классификатора

$$Q(a, X) = \frac{1}{l}\sum_{i=1}^{l}[a(x_i) \neq y_i] \to \min_a$$ - доля ошибок

Обучение линейного классификатора

$Q(a, X) = \frac{1}{l}\sum_{i=1}^{l}[a(x_i) \neq y_i] \to \min_a$ - доля ошибок

$$\frac{1}{l}\sum_{i=1}^{l}[sign(< w, x_i >) \neq y_i] \to \min_w$$

Обучение линейного классификатора

$Q(a, X) = \frac{1}{l}\sum_{i=1}^{l}[a(x_i) \neq y_i] \rightarrow \min_a$ - доля ошибок

$$\frac{1}{l}\sum_{i=1}^{l}[sign(< w, x_i >) \neq y_i] \rightarrow \min_w$$

$$\frac{1}{l}\sum_{i=1}^{l}[y_i < w, x_i > \quad < 0] \rightarrow \min_w$$

Обучение линейного классификатора

$Q(a, X) = \frac{1}{l}\sum_{i=1}^{l}[a(x_i) \neq y_i] \rightarrow \min_{a}$ - доля ошибок

$$\frac{1}{l}\sum_{i=1}^{l}[sign(< w, x_i >) \neq y_i] \rightarrow \min_{w}$$

$$\frac{1}{l}\sum_{i=1}^{l}[y_i < w, x_i > \ < 0] \rightarrow \min_{w}$$

$y_i < w, x_i > \ > 0$ – правильный ответ

$y_i < w, x_i > \ < 0$ – неправильный ответ

Обучение линейного классификатора

$Q(a,X) = \frac{1}{l}\sum_{i=1}^{l}[a(x_i) \neq y_i] \to \min_a$ - доля ошибок

$$\frac{1}{l}\sum_{i=1}^{l}[sign(< w, x_i >) \neq y_i] \to \min_w$$

$$\frac{1}{l}\sum_{i=1}^{l}[y_i < w, x_i > \ < 0] \to \min_w$$

$y_i < w, x_i > \ > 0$ – правильный ответ

$y_i < w, x_i > \ < 0$ – неправильный ответ

$M_i = y_i < w, x_i >$ - отступ (margin)

Обучение линейного классификатора

$M_i = \; y_i < w, x_i >$ - отступ (margin)

$|M_i|$ - расстояние от объекта i до разделяющей гиперплоскости

$|M_i|$ - уверенность модели

$M_i \ll 0$ – модель сильно уверенна, но ошибается.

Скорее всего это выброс.

Обучение линейного классификатора

$$\frac{1}{l}\sum_{i=1}^{l}[y_i < w, x_i > \quad < 0] \to \min_{w}$$

Обучение линейного классификатора

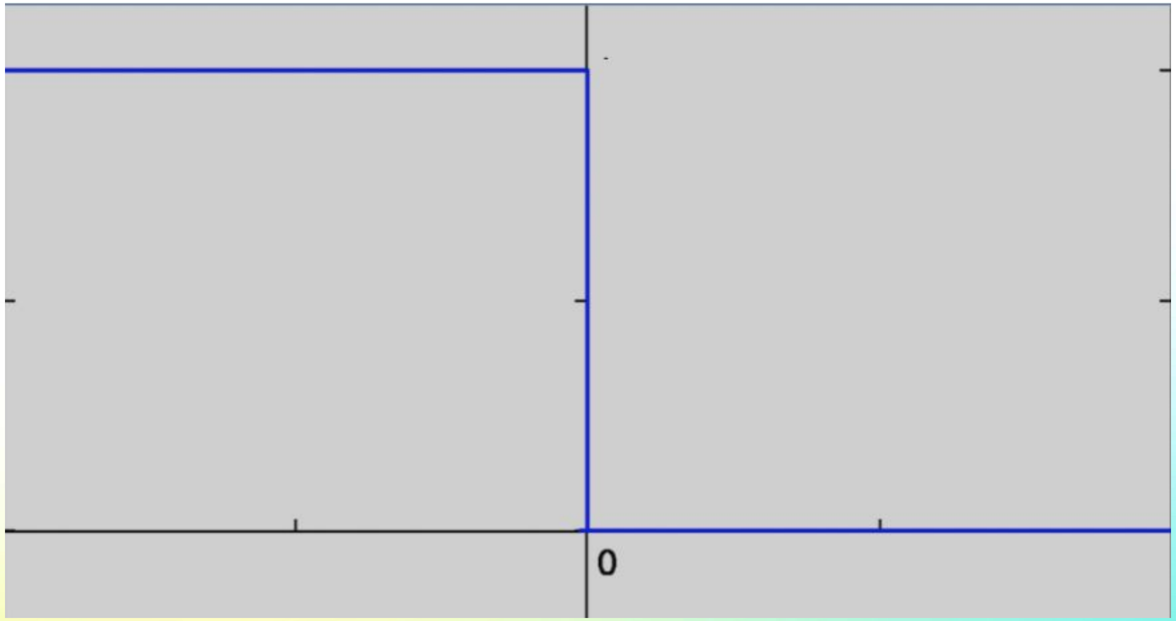
$$\frac{1}{l}\sum_{i=1}^{l}[y_i < w, x_i > \; < 0] \to \min_{w}$$

$L(M) = [M < 0]$ – пороговая функция потерь

Обучение линейного классификатора

$$\frac{1}{l}\sum_{i=1}^{l}[y_i < w, x_i > \;\; < 0] \to \min_{w}$$

$L(M) = [M < 0]$ – пороговая функция потерь



Обучение линейного классификатора

$$0 \leq \frac{1}{l}\sum_{i=1}^{l}[y_i < w, x_i > \;\; < 0] \leq \frac{1}{l}\sum_{i=1}^{l}\tilde{L}(y_i < w, x_i >) \to \min_{w}$$

Обучение линейного классификатора

$$0 \le \frac{1}{l}\sum_{i=1}^{l}[y_i < w, x_i > \; < 0] \le \frac{1}{l}\sum_{i=1}^{l}\tilde{L}(y_i < w, x_i >) \;\to \min_{w}$$

$\tilde{L}$ - дифференцируемая функция

Обучение линейного классификатора

$$0 \le \frac{1}{l}\sum_{i=1}^{l}[y_i < w, x_i > \ < 0] \le \frac{1}{l}\sum_{i=1}^{l}\tilde{L}(y_i < w, x_i >) \rightarrow \min_{w}$$

$\tilde{L}$ - дифференцируемая функция

$\tilde{L}(M) = \ e^{-M}$ - экспоненциальная функция потерь

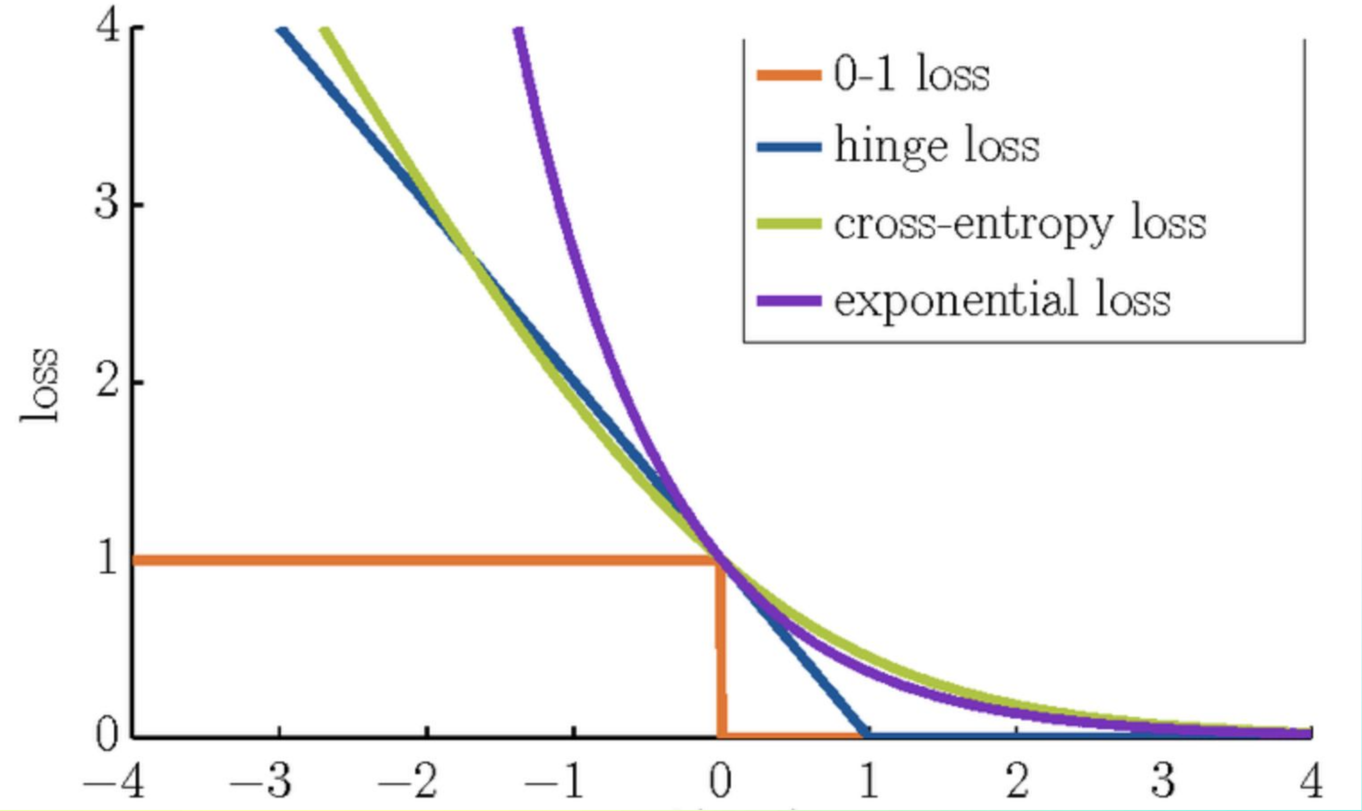
$\tilde{L}(M) = \log(1 + \ e^{-M})$ - логарифмическая функция потерь

$\tilde{L}(M) = \ \frac{2}{1+ e^{M}}$ - сигмоидная функция потерь

$\tilde{L}(M) = \ (1 - M)^2$ - квадратичная функция потерь

$\tilde{L}(M) = \max(0, 1 - M)$ – кусочно-линейная функция потерь

Обучение линейного классификатора



Измерение ошибки в задачах классификации. Метрики качества

Доля правильных ответов(аккуратность)

$$accuracy(a, X) = \frac{1}{l}\sum_{i=1}^{l}[a(x_i) = y_i]$$

Измерение ошибки в задачах классификации. Метрики качества

Доля правильных ответов(аккуратность)

$$accuracy(a, X) = \frac{1}{l}\sum_{i=1}^{l}[a(x_i) = y_i]$$

95 объектов – 1 класс

5 объектов – -1 класс

Измерение ошибки в задачах классификации. Метрики качества

Доля правильных ответов(аккуратность)

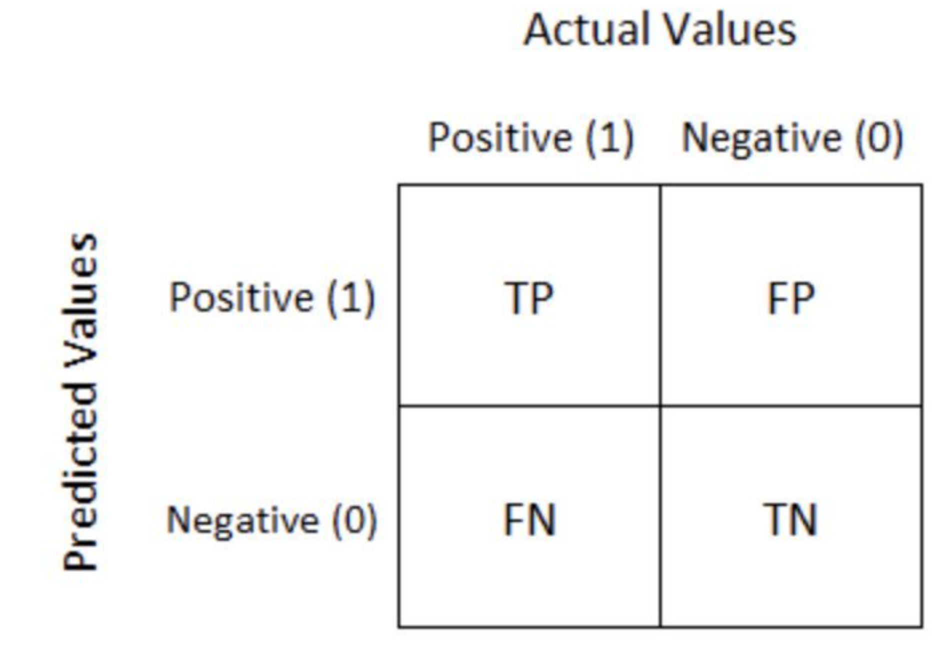
$$accuracy(a, X) = \frac{1}{l}\sum_{i=1}^{l}[a(x_i) = y_i]$$

95 объектов – 1 класс

5 объектов – -1 класс

accuracy = 0.95

Измерение ошибки в задачах классификации. Метрики качества



Измерение ошибки в задачах классификации. Метрики качества

Точность

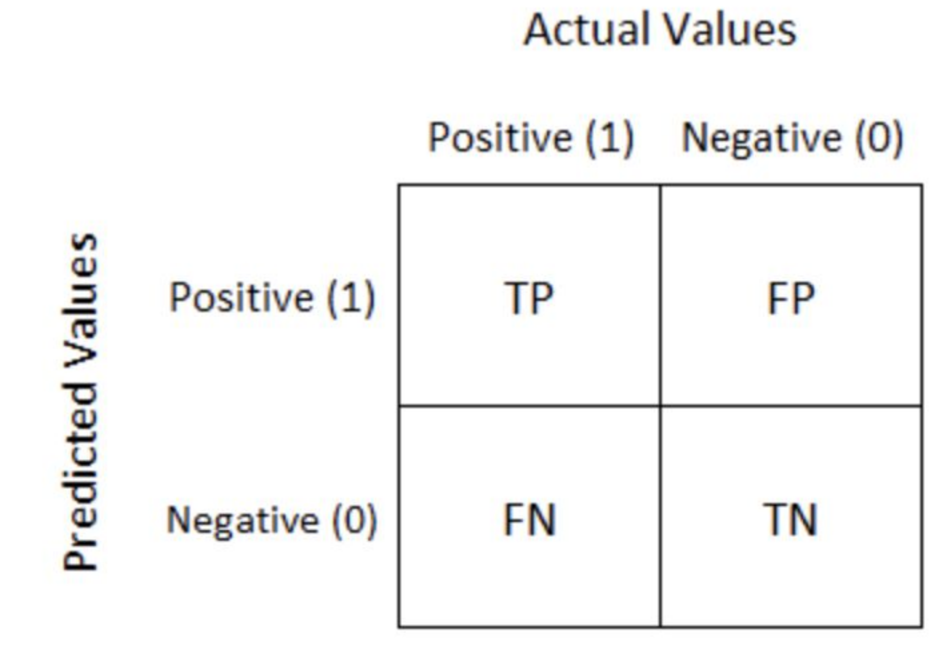
$$precision = \frac{TP}{TP + FP}$$

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

Измерение ошибки в задачах классификации. Метрики качества

Полнота

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



Измерение ошибки в задачах классификации. Метрики качества

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

Измерение ошибки в задачах классификации. Метрики качества

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

precision = 0.1

recall = 1

Измерение ошибки в задачах классификации. Метрики качества

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

precision = 0.1

recall = 1

A = 0.55

Измерение ошибки в задачах классификации. Метрики качества

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

precision = 0.1

recall = 1

A = 0.55

precision=0.55

recall=0.55

Измерение ошибки в задачах классификации. Метрики качества

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

precision = 0.1

recall = 1

A = 0.55

precision=0.55

recall=0.55

A = 0.55

Измерение ошибки в задачах классификации. Метрики качества

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

Измерение ошибки в задачах классификации. Метрики качества

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

precision = 0.4

recall = 0.9

Измерение ошибки в задачах классификации. Метрики качества

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

precision = 0.4

recall = 0.9

M = 0.4

Измерение ошибки в задачах классификации. Метрики качества

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

precision = 0.4

recall = 0.9

M = 0.4

precision=0.4

recall=1

Измерение ошибки в задачах классификации. Метрики качества

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

precision = 0.4

recall = 0.9

M = 0.4

precision=0.4

recall=1

M = 0.4

Измерение ошибки в задачах классификации. Метрики качества

F1-мера

$$F1 = \frac{2 * precision * recall}{precision + recall}$$

Измерение ошибки в задачах классификации. Метрики качества

F1-мера

$$\mathrm{F1} = \frac{2 * precision * recall}{precision + recall}$$

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

Измерение ошибки в задачах классификации. Метрики качества

F1-мера

$$\text{F1} = \frac{2 * precision * recall}{precision + recall}$$

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

$\beta = 0.5$ – важнее точность

$\beta = 2$ – важнее полнота

Измерение ошибки в задачах классификации. Метрики качества

$$\mathrm{G} = \sqrt{precision * recall}$$

Измерение ошибки в задачах классификации. Метрики качества

$$G = \sqrt{precision * recall}$$

precision = 0.9

recall = 0.1

Измерение ошибки в задачах классификации. Метрики качества

$$G = \sqrt{precision * recall}$$

precision = 0.9

recall = 0.1

G = 0.3

Измерение ошибки в задачах классификации. Метрики качества

$$G = \sqrt{precision * recall}$$

precision = 0.9

recall = 0.1

G = 0.3

F1 = 0.18

Метрики качества ранжирования

$a(x) = sign < w, x >$

Метрики качества ранжирования

$a(x) = sign < w, x >$

$a(x) \in \{-1, 1\}$

Метрики качества ранжирования

$a(x) = sign < w, x >$

$a(x) \in \{-1, 1\}$

$a(x) = sign(< w, x > -t)$

Метрики качества ранжирования

$$a(x) = sign < w, x >$$

$$a(x) \in \{-1, 1\}$$

$$a(x) = sign(< w, x > -t)$$

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

Метрики качества ранжирования

<w,x>
50
40
35
20
10
5
1
-1
-5
-20
-30

Метрики качества ранжирования

<w,x>	y
50	+1
40	+1
35	+1
20	+1
10	+1
5	+1
1	+1
-1	-1
-5	-1
-20	-1
-30	-1

Метрики качества ранжирования

<w,x>	y	y
50	+1	+1
40	+1	+1
35	+1	-1
20	+1	+1
10	+1	-1
5	+1	+1
1	+1	+1
-1	-1	-1
-5	-1	+1
-20	-1	-1
-30	-1	-1

Метрики качества ранжирования

$a(x) = [< w, x > \; > t]$

Метрики качества ранжирования

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

$b(x) = < w, x >$ - показатель уверенности

Метрики качества ранжирования

$a(x) = [< w, x > \;\; > t]$

$b(x) = < w, x >$ - показатель уверенности

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

Метрики качества ранжирования

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

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

Метрики качества ранжирования

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

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

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

Метрики качества ранжирования

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

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

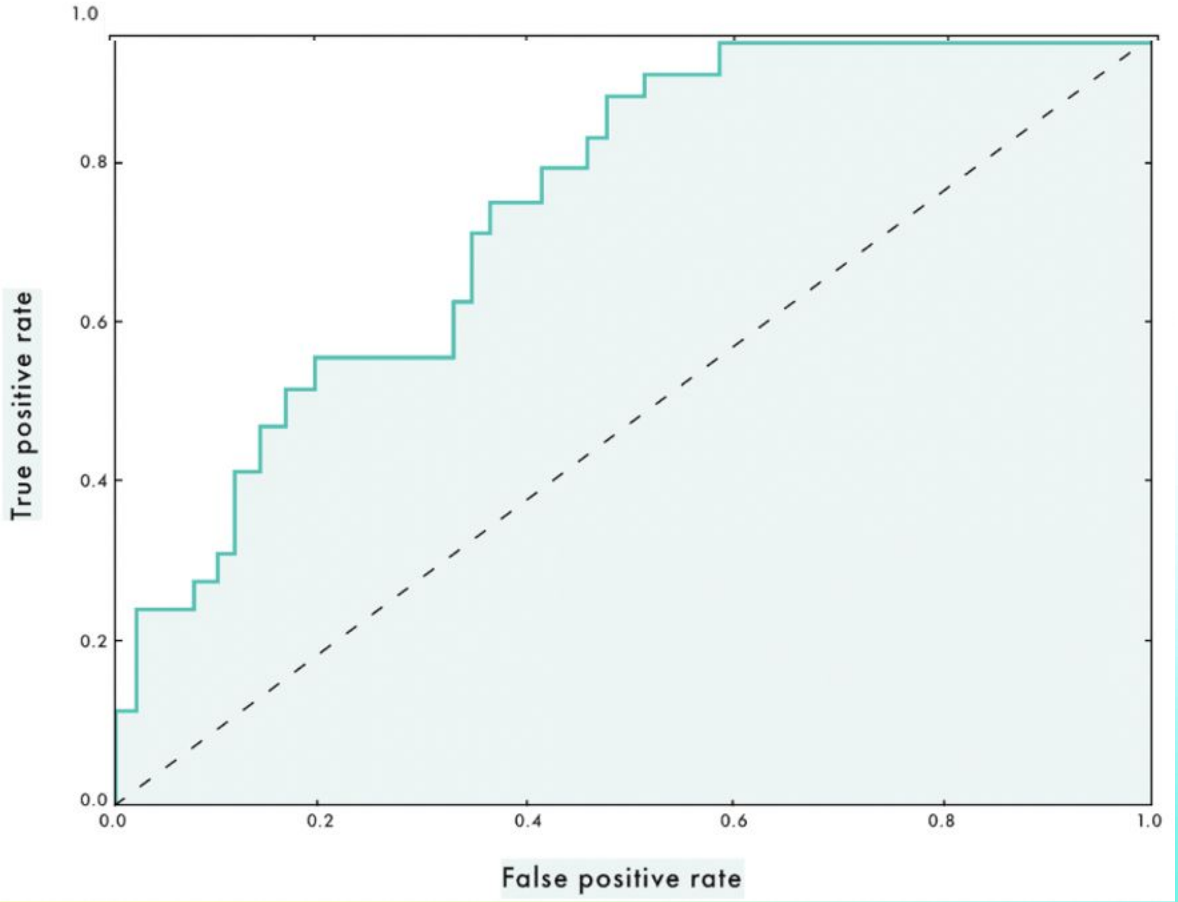
$$b(x_1) \leq b(x_2) \leq b(x_3) \leq \cdots \leq b(x_l)$$

Метрики качества ранжирования

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

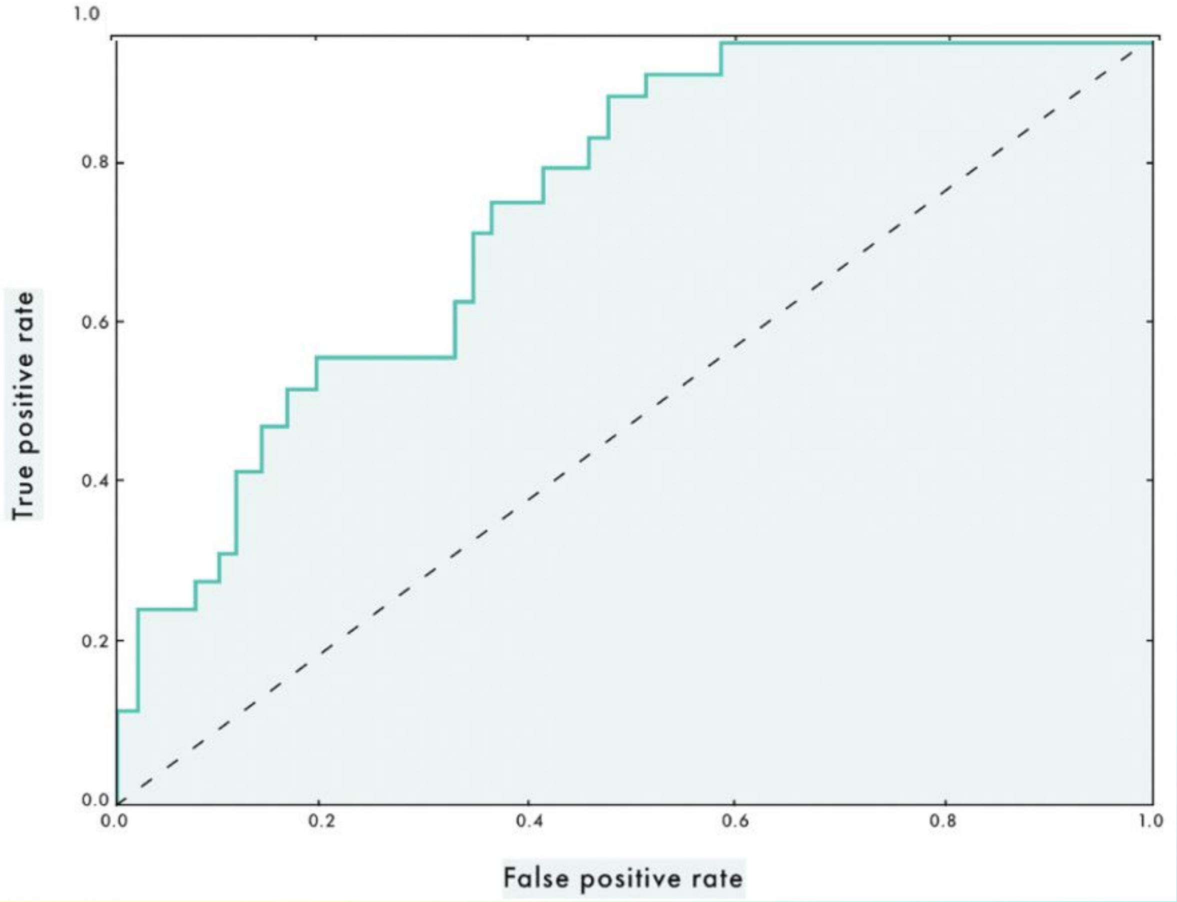
$$TPR = \frac{TP}{TP + FN}$$

$$b(x_1) \leq b(x_2) \leq b(x_3) \leq \cdots \leq b(x_l)$$



Метрики качества ранжирования

ROC-AUC – площадь под TPR-FPR кривой



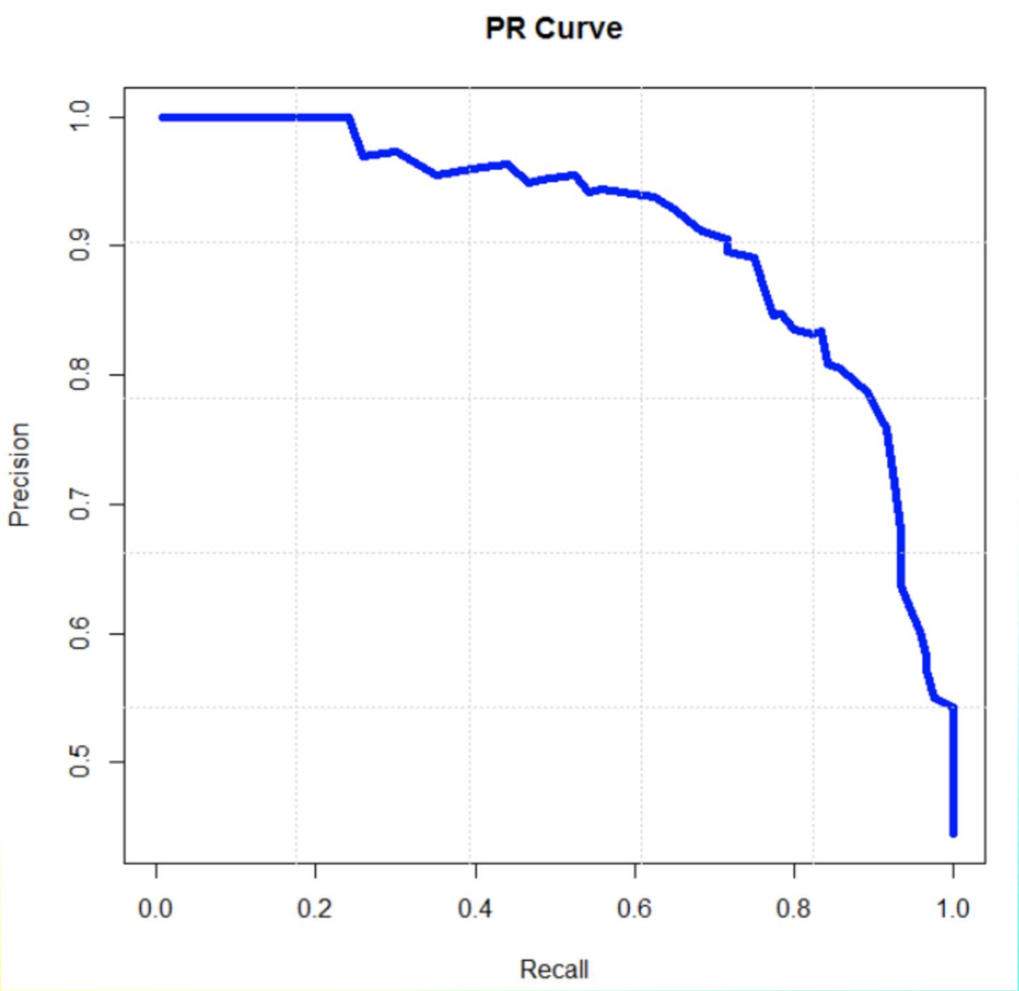
Метрики качества ранжирования

Индекс Джини

GINI = 2* ROC-AUC - 1

Метрики качества ранжирования

Precision-Recall curve



Метрики качества ранжирования

$100 \in \{+1\}$

$1\,000\,000 \in \{-1\}$

t:

$95 \in \{+1\}$

$50\,000 \in \{-1\}$

TPR = 0.95

FPR = 0.05

ROC-AUC = 0.95

PRC-AUC = 0.001

Логистическая регрессия

$b(x) \in [0,1]$

Логистическая регрессия

$b(x) \in [0,1]$

$b(x) = 0.9$ (модель уверенна на 90% в том что объект принадлежит положительному классу)

Логистическая регрессия

$b(x) \in [0,1]$

$b(x) = 0.9$ (модель уверенна на 90% в том что объект принадлежит положительному классу)

$b(x) = p(y = +1 | x)$

Логистическая регрессия

user_id	product_id	target
1	10	0
1	10	0
1	10	0
1	10	0
1	10	0
1	10	0
1	10	0
1	10	0
1	10	0
1	10	1

$$b(x) = 0.1$$

Логистическая регрессия

c(x)	p(y= +1| x)	c(x)*p(y= +1| x)
500	0,5	250
1000	0,3	300
2000	0,2	400
5000	0,01	50
10000	0,001	10

Логистическая регрессия

$\{x_1, x_2, \dots, x_n\}: b(x_1) = b(x_2) = \dots = b(x_n) = b = const$

$\{y_1, y_2, \dots, y_n\}$

Логистическая регрессия

$\{x_1, x_2, \dots, x_n\}: b(x_1) = b(x_2) = \dots = b(x_n) = b = const$
$\{y_1, y_2, \dots, y_n\}$

$$\underset{b}{argmin}\ \frac{1}{n}\sum_{i=1}^{n} L(y_i, b) = p(y = +1 |\, x)$$

при $n \to \infty$

Логистическая регрессия

$\{x_1, x_2, \dots, x_n\}: b(x_1) = b(x_2) = \dots = b(x_n) = b = const$
$\{y_1, y_2, \dots, y_n\}$

$$arg\min_b \frac{1}{n}\sum_{i=1}^{n} L(y_i, b) = \frac{1}{n}\sum_{i=1}^{n} [y_i = +1]$$

при $n \to \infty$

$arg\min_b E(L(y, b) | x) = p(y = +1 | x)$ – условие корректного оценивания вероятности

Логистическая регрессия

$\{x_1, x_2, \dots, x_n\}: b(x_1) = b(x_2) = \dots = b(x_n) = b = const$
$\{y_1, y_2, \dots, y_n\}$

$$\underset{b}{argmin} \ \frac{1}{n}\sum_{i=1}^{n} L(y_i, b) = p(y = +1 | x)$$

при $n \to \infty$

$\underset{b}{argmin} \ E(L(y, b) | x) = p(y = +1 | x)$ – условие корректного оценивания вероятности

$$\mathrm{E} = p(y = +1 | x) * \mathrm{L}(+1, \mathrm{b}) + p(y = -1 | x) * \mathrm{L}(-1, \mathrm{b})$$

Логистическая регрессия

$b \in [0,1]$

Логистическая регрессия

$b \in [0,1]$

$< w, x > \in R$

Логистическая регрессия

$b \in [0,1]$

$< w, x > \in R$

$b(x) = \sigma(< w, x >)$

$\sigma(z) = \dfrac{1}{1 + e^{-z}}$

Логистическая регрессия

$b \in [0,1]$

$< w, x > \in R$

$b(x) = \sigma(< w, x >)$

$\sigma(z) = \dfrac{1}{1 + e^{-z}}$

$$\frac{1}{l}\sum_{i=1}^{l}(\sigma(< w, x_i >) - y_i)^2 \rightarrow \min_{w}$$

Логистическая регрессия

$\prod_{i=1}^{l} b(x_i)^{[y_i = +1]}(1 - b(x_i))^{[y_i = -1]} \to max$ - правдоподобие

Логистическая регрессия

$\prod_{i=1}^{l} b(x_i)^{[y_i = +1]}(1 - b(x_i))^{[y_i = -1]} \to max$ - правдоподобие

$$-\sum_{i=1}^{l}([y_i = \ +1] * \log\big(b(x_i)\big) + [y_i = -1] * log(1 - b(x_i))) \to \min_{b}$$

Логистическая регрессия

$\prod_{i=1}^{l} b(x_i)^{[y_i = +1]}(1 - b(x_i))^{[y_i = -1]} \rightarrow max$ - правдоподобие

$$-\sum_{i=1}^{l}([y_i = +1] * \log\big(b(x_i)\big) + [y_i = -1] * log(1 - b(x_i))) \rightarrow \min_b$$

$L(y, b) = -[y = +1] * \log(b) - [y = -1] * log(1 - b))$ – log-loss

Логистическая регрессия

$$E(L(y,b)|x) = -p(y=+1|\ x) * \log(b) - p(y=-1|\ x) * log(1-b))$$

Логистическая регрессия

$$E(L(y,b)|x) = -p(y=+1|\,x) * \log(b) - p(y=-1|\,x) * log(1-b))$$

$$\frac{\partial}{\partial b} = -\frac{p(y=+1|\,x)}{b} + \frac{1-p(y=+1|\,x)}{1-b} = 0$$

Логистическая регрессия

$$E(L(y,b)|x) = -p(y=+1|\,x) * \log(b) - p(y=-1|\,x) * log(1-b))$$

$$\frac{\partial}{\partial b} = -\frac{p(y=+1|\,x)}{b} + \frac{1-p(y=+1|\,x)}{1-b} = 0$$

$$b = \, p(y=+1|\,x)$$

Логистическая регрессия

$$-\sum_{i=1}^{l}\left([y_i = \ +1] * \log\left(\frac{1}{1+e^{-<w,x>}}\right) + [y_i = -1] * log\left(1 - \frac{1}{1+e^{-<w,x>}}\right)\right)$$

Логистическая регрессия

$$-\sum_{i=1}^{l}\left(\left[y_i = \ +1\right] * \log\left(\frac{1}{1+e^{-<w,x>}}\right) + \left[y_i = -1\right] * log\left(1 - \frac{1}{1+e^{-<w,x>}}\right)\right) =$$

$$= -\sum_{i=1}^{l}\left(\left[y_i = \ +1\right] * \log\left(\frac{1}{1+e^{-<w,x>}}\right) + \left[y_i = -1\right] * log\left(\frac{1}{1+e^{<w,x>}}\right)\right)$$

Логистическая регрессия

$$-\sum_{i=1}^{l}\left([y_i = \ +1] * \log\left(\frac{1}{1+e^{-<w,x>}}\right) + [y_i = -1] * log\left(1 - \frac{1}{1+e^{-<w,x>}}\right)\right) =$$

$$= -\sum_{i=1}^{l}\left([y_i = \ +1] * \log\left(\frac{1}{1+e^{-<w,x>}}\right) + [y_i = -1] * log\left(\frac{1}{1+e^{<w,x>}}\right)\right) =$$

$$= -\sum_{i=1}^{l} log\left(\frac{1}{1+e^{-y_i<w,x>}}\right)$$

Логистическая регрессия

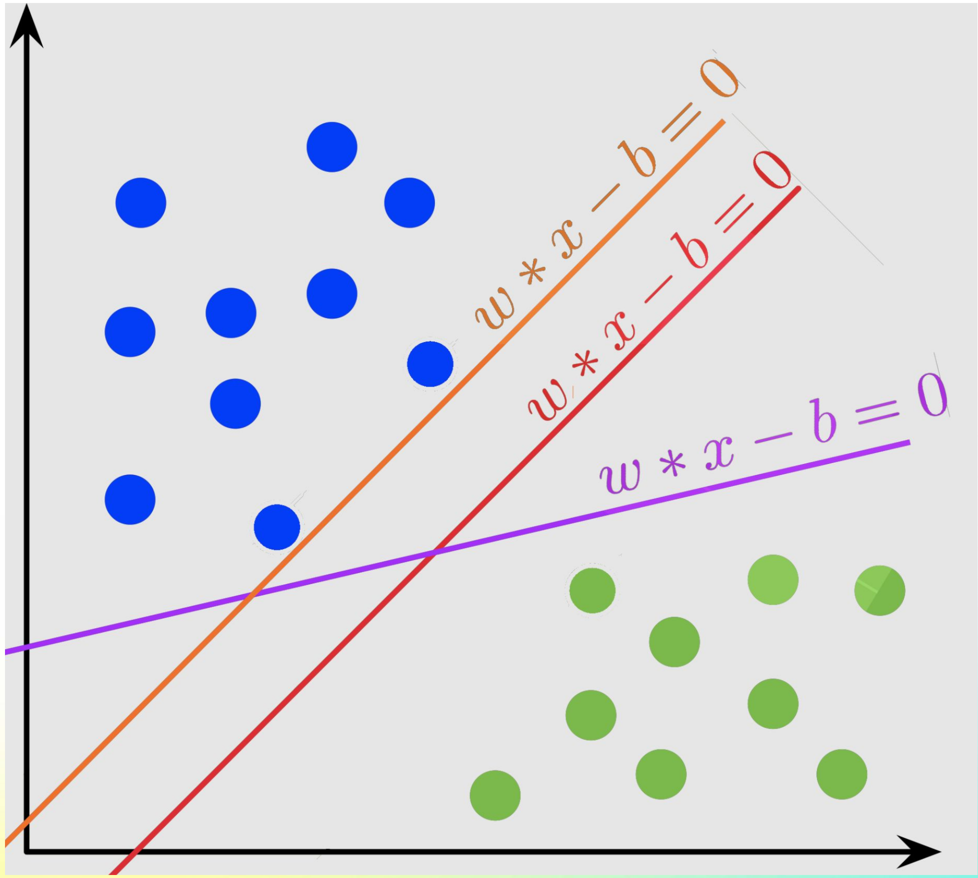
$$-\sum_{i=1}^{l}\left(\left[y_i = \ +1\right] * \log\left(\frac{1}{1+e^{-<w,x>}}\right) + \left[y_i = -1\right] * log\left(1 - \frac{1}{1+e^{-<w,x>}}\right)\right) =$$

$$= -\sum_{i=1}^{l}\left(\left[y_i = \ +1\right] * \log\left(\frac{1}{1+e^{-<w,x>}}\right) + \left[y_i = -1\right] * log\left(\frac{1}{1+e^{<w,x>}}\right)\right) =$$

$$= -\sum_{i=1}^{l} log\left(\frac{1}{1+e^{-y_i<w,x>}}\right) = \sum_{i=1}^{l} log(1 + e^{-y_i<w,x>}) \to \min_{w}$$

Метод опорных векторов

Support vector machine (SVM)
$L(M) = \max(0, 1 - M)$



Метод опорных векторов

1) Линейно разделимый случай

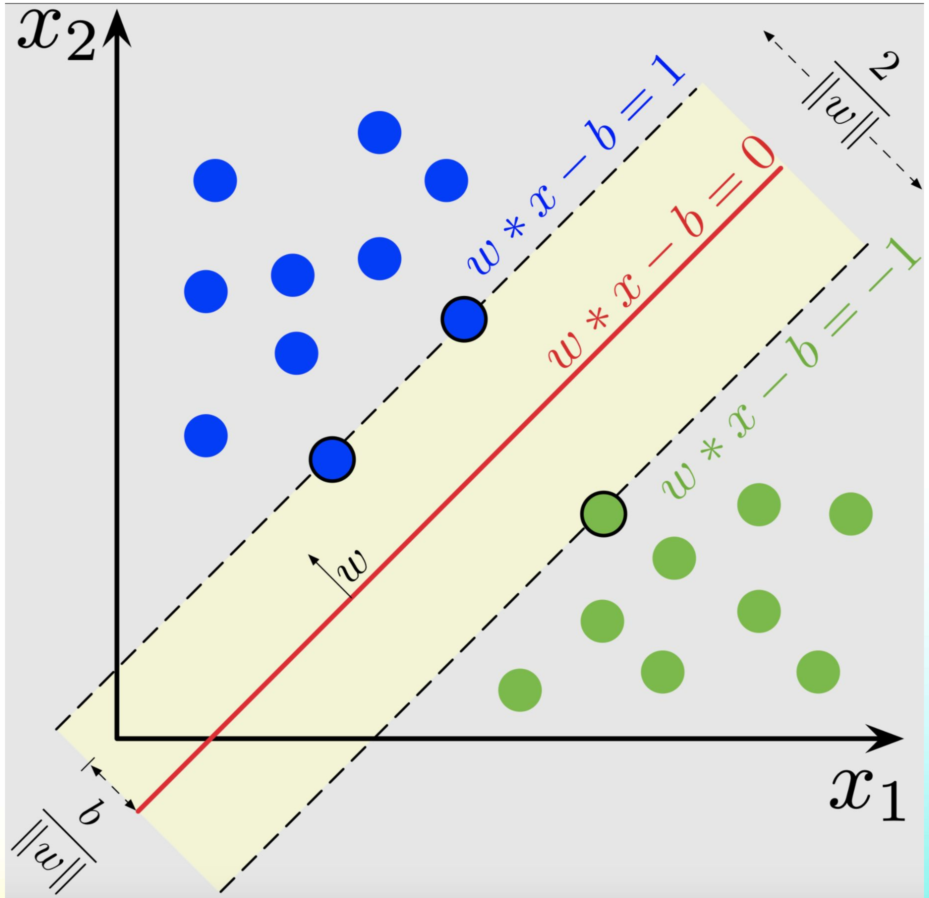
$a(x) = sign(< w, x > +w_0)$

$\min\limits_{x} |< w, x > +w_0| = 1$

$\rho(x, a) = \frac{|<w,x>+w_0|}{\|w\|_2}$ - расстояние от объекта x до разделяющей гиперплоскости

$\min\limits_{x} \frac{|<w,x>+w_0|}{\|w\|_2} = \frac{1}{\|w\|_2}$ - минимальное расстояние от разделяющей гиперплоскости до объектов выборки

$\frac{2}{\|w\|_2}$ - ширина разделяющей полоса(отступ классификатора)



Метод опорных векторов

$$\frac{1}{\|w\|_2^2} \to \max_{w}$$

Метод опорных векторов

$$\frac{1}{\|w\|_2^2} \to \max_w$$

$$\|w\|_2^2 \to \min_w$$

Метод опорных векторов

$$\frac{1}{\|w\|_2^2} \to \max_w$$

$$\|w\|_2^2 \to \min_w$$

$$y_i\,(< w, x > + w_0) \geq 0$$

Метод опорных векторов

$$\frac{1}{\|w\|_2^2} \to \max_w$$

$$\|w\|_2^2 \to \min_w$$

$$y_i\ (< w, x > + w_0) \geq 0$$

$$|< w, x + b >| \geq 1$$

Метод опорных векторов

$$\frac{1}{\|w\|_2^2} \to \max_w$$

$$\|w\|_2^2 \to \min_w$$

$$y_i\,(< w, x > + w_0) \geq 0$$

$$|< w, x + b >| \geq 1$$

$$\begin{cases} \|w\|_2^2 \to \min\limits_w \\ y_i\,(< w, x > + w_0) \geq 1 \end{cases}$$

Метод опорных векторов

2) Линейно неразделимый случай

$$\begin{cases} \|w\|_2^2 + C\dfrac{1}{l}\sum\limits_{i=1}^{l}\xi_i \to \min\limits_{w} \\ y_i\,(< w, x > +w_0) \geq 1\text{-}\xi_i \\ \xi_i \geq 0 \end{cases}$$

C = гиперпараметр

Метод опорных векторов

3) Безусловный случай

$$\begin{cases} \xi_i \geq 1 - y_i\,(< w, x > + w_0) \\ \xi_i \geq 0 \end{cases}$$

Метод опорных векторов

3) Безусловный случай

$$\begin{cases} \xi_i \geq 1 - y_i\,(< w, x > +w_0) \\ \xi_i \geq 0 \end{cases}$$

$$\xi_i = \max(0, 1 - y_i\,(< w, x > +w_0))$$

Метод опорных векторов

3) Безусловный случай

$$\begin{cases} \xi_i \geq 1 - y_i\,(< w, x > + w_0) \\ \xi_i \geq 0 \end{cases}$$

$$\xi_i = \max(0, 1 - y_i\,(< w, x > + w_0))$$

$$\|w\|_2^2 + C\frac{1}{l}\sum_{i=1}^{l} \max(0, 1 - y_i\,(< w, x > + w_0)) \to \min_w$$

Метод опорных векторов

3) Безусловный случай

$$\begin{cases} \xi_i \geq 1 - y_i \,(< w, x > + w_0) \\ \xi_i \geq 0 \end{cases}$$

$$\xi_i = \max(0, 1 - y_i \,(< w, x > + w_0))$$

$$\|w\|_2^2 + C \frac{1}{l} \sum_{i=1}^{l} \max(0, 1 - y_i \,(< w, x > + w_0)) \to \min_w$$

$$\frac{1}{l} \sum_{i=1}^{l} \max(0, 1 - M_i) + \frac{1}{C} \|w\|_2^2 \to \min_w$$

Метод опорных векторов

3) Безусловный случай

$$\begin{cases} \xi_i \geq 1 - y_i\,(< w, x > + w_0) \\ \xi_i \geq 0 \end{cases}$$

$$\xi_i = \max(0, 1 - y_i\,(< w, x > + w_0))$$

$$\|w\|_2^2 + C\frac{1}{l}\sum_{i=1}^{l}\max(0, 1 - y_i\,(< w, x > + w_0)) \to \min_w$$

$$\frac{1}{l}\sum_{i=1}^{l}\max(0, 1 - M_i) + \frac{1}{C}\|w\|_2^2 \to \min_w$$

$$[\mathrm{M} < 0] \leq \max(0, 1 - M)$$

Метод опорных векторов

3) Безусловный случай

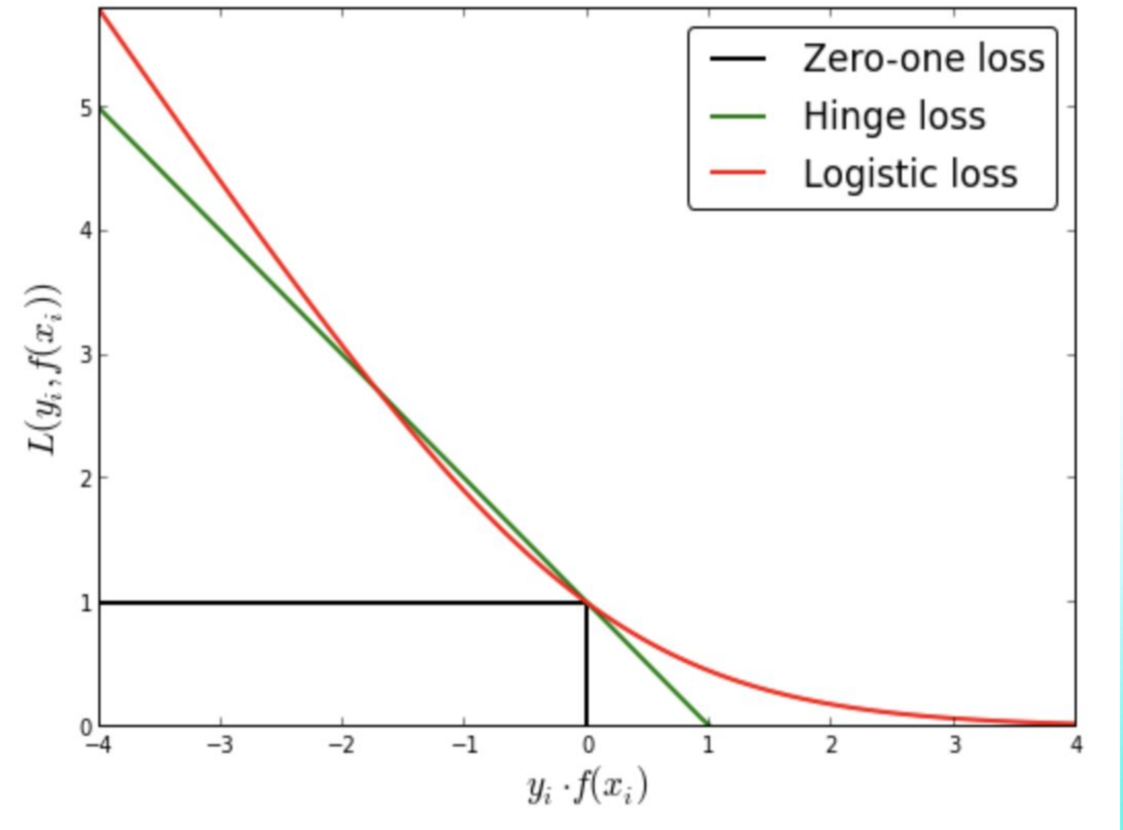
$$\begin{cases} \xi_i \geq 1 - y_i\,(< w, x > + w_0) \\ \xi_i \geq 0 \end{cases}$$

$$\xi_i = \max(0, 1 - y_i\,(< w, x > + w_0))$$

$$\|w\|_2^2 + C\,\frac{1}{l}\sum_{i=1}^{l} \max(0, 1 - y_i\,(< w, x > + w_0)) \to \min_w$$

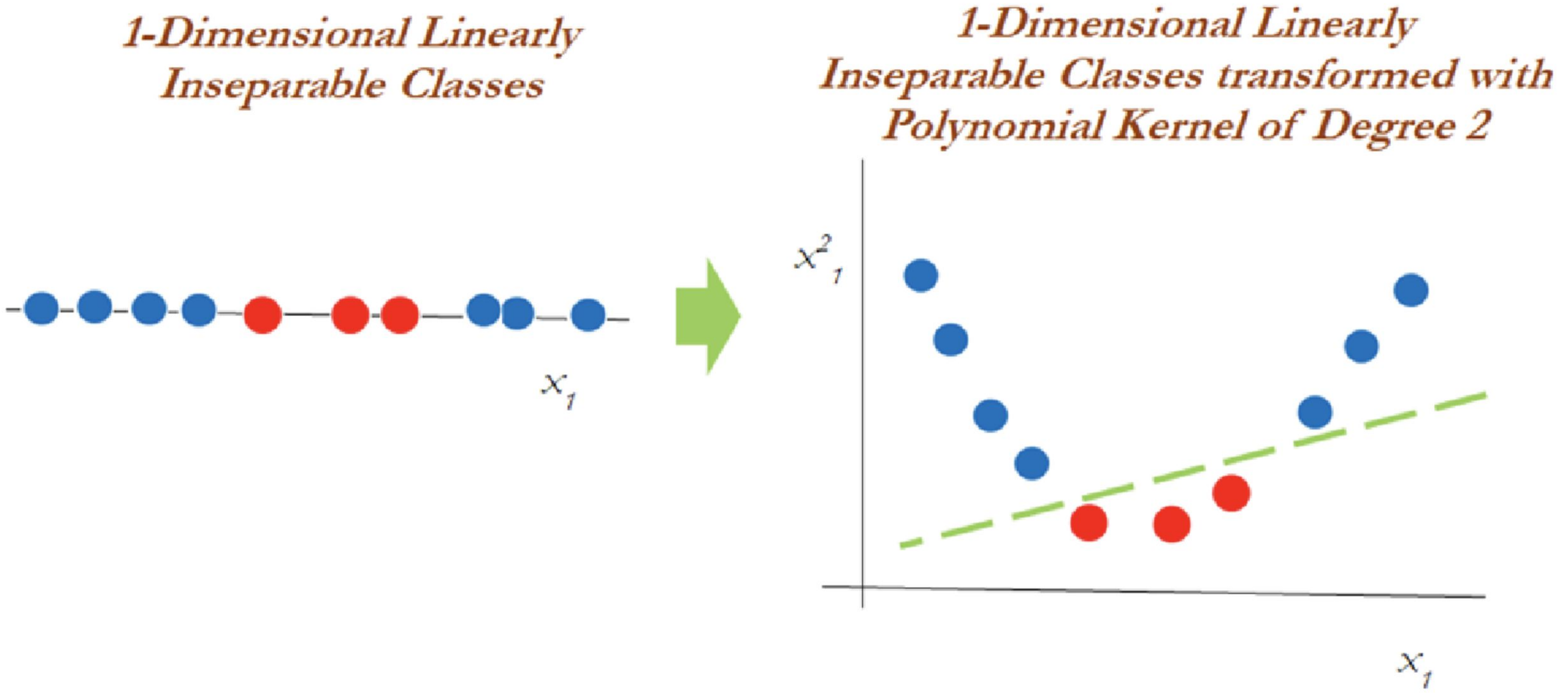
$$\frac{1}{l}\sum_{i=1}^{l} \max(0, 1 - M_i) + \frac{1}{C}\|w\|_2^2 \to \min_w$$

$$[\mathrm{M} < 0] \leq \max(0, 1 - M)$$



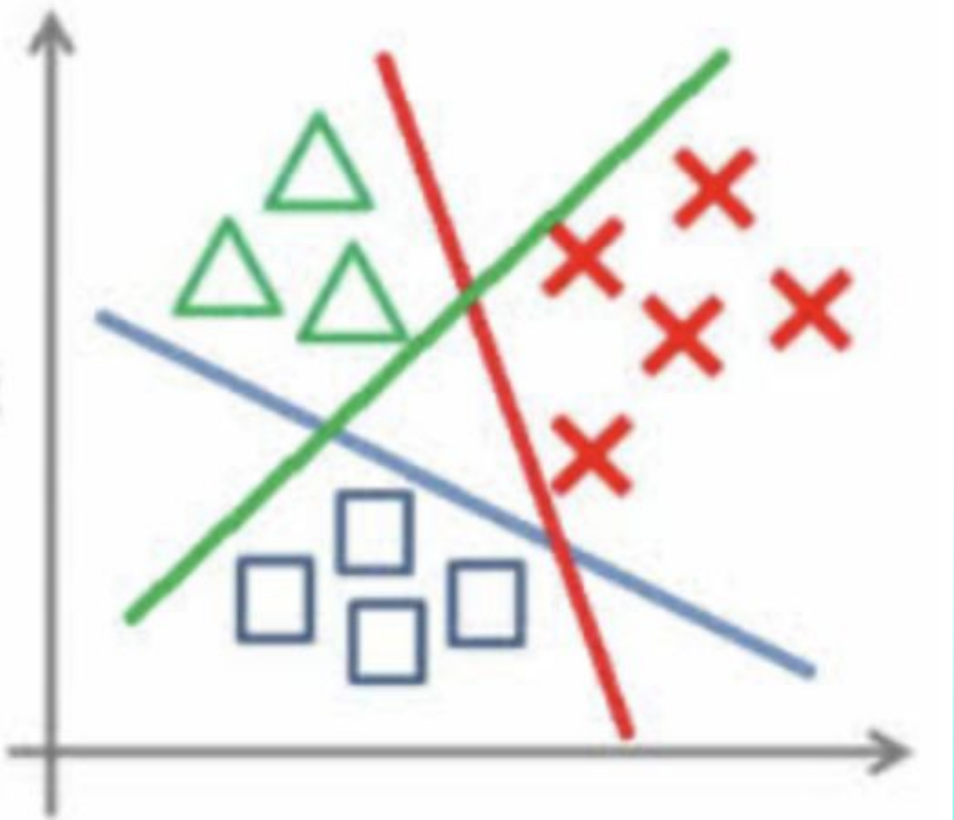
Метод опорных векторов

ядерный метод (kernel trick)



Многоклассовая классификация

1) Один против всех (one-vs-all, one-vs-rest)

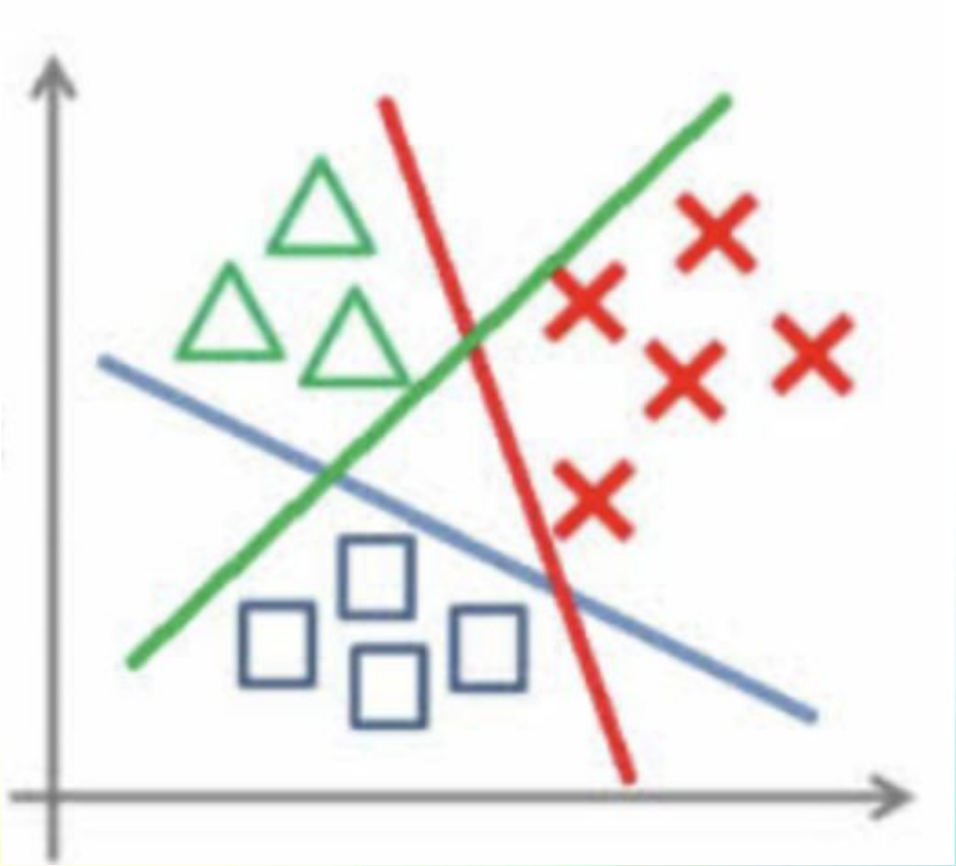


Многоклассовая классификация

1) Один против всех (one-vs-all, one-vs-rest)

$$X_k = \{(x_i, 2[y_i = k] - 1)\}_{i=1}^{l}$$

$$a_x(x) = sign(b_k(x))$$



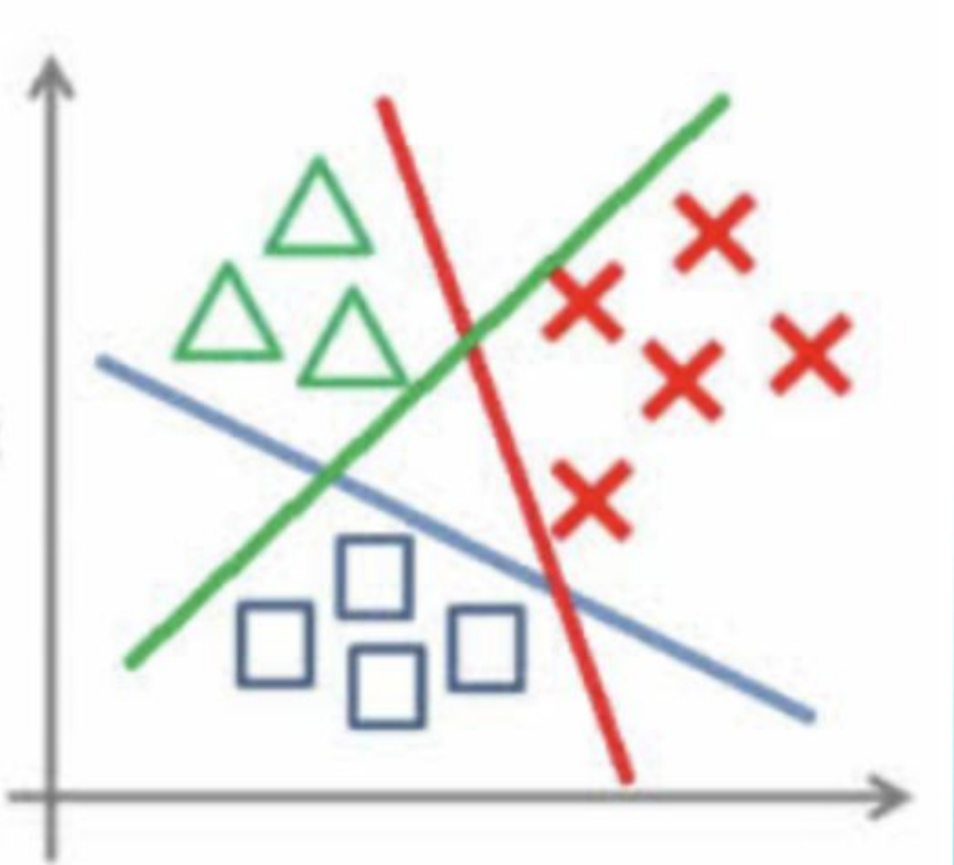
Многоклассовая классификация

1) Один против всех (one-vs-all, one-vs-rest)

$$X_k = \{(x_i, 2[y_i = k] - 1)\}_{i=1}^{l}$$

$$a_x(x) = sign(b_k(x))$$

$$b_k(x) =< w_k, x > +w_{0k}$$



Многоклассовая классификация

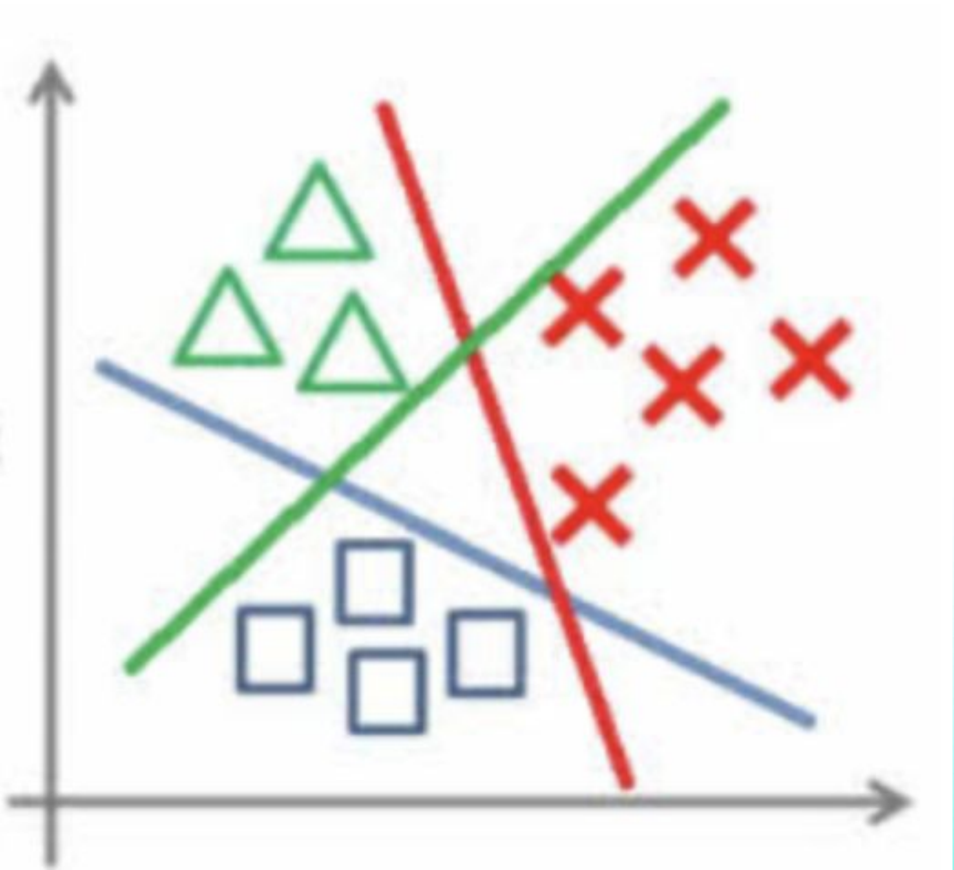
1) Один против всех (one-vs-all, one-vs-rest)

$$X_k = \{(x_i, 2[y_i = k] - 1)\}_{i=1}^{l}$$

$$a_x(x) = sign(b_k(x))$$

$$b_k(x) = < w_k, x > + w_{0k}$$

$$a(x) = \underset{w}{\operatorname{argmax}}(b_k(x))$$



Многоклассовая классификация

1) Один против всех (one-vs-all, one-vs-rest)

$$X_k = \{(x_i, 2[y_i = k] - 1)\}_{i=1}^{l}$$

$$a_x(x) = sign(b_k(x))$$

$$b_k(x) =< w_k, x > +w_{0k}$$

$$a(x) = \underset{w}{\mathrm{argmax}}(b_k(x))$$

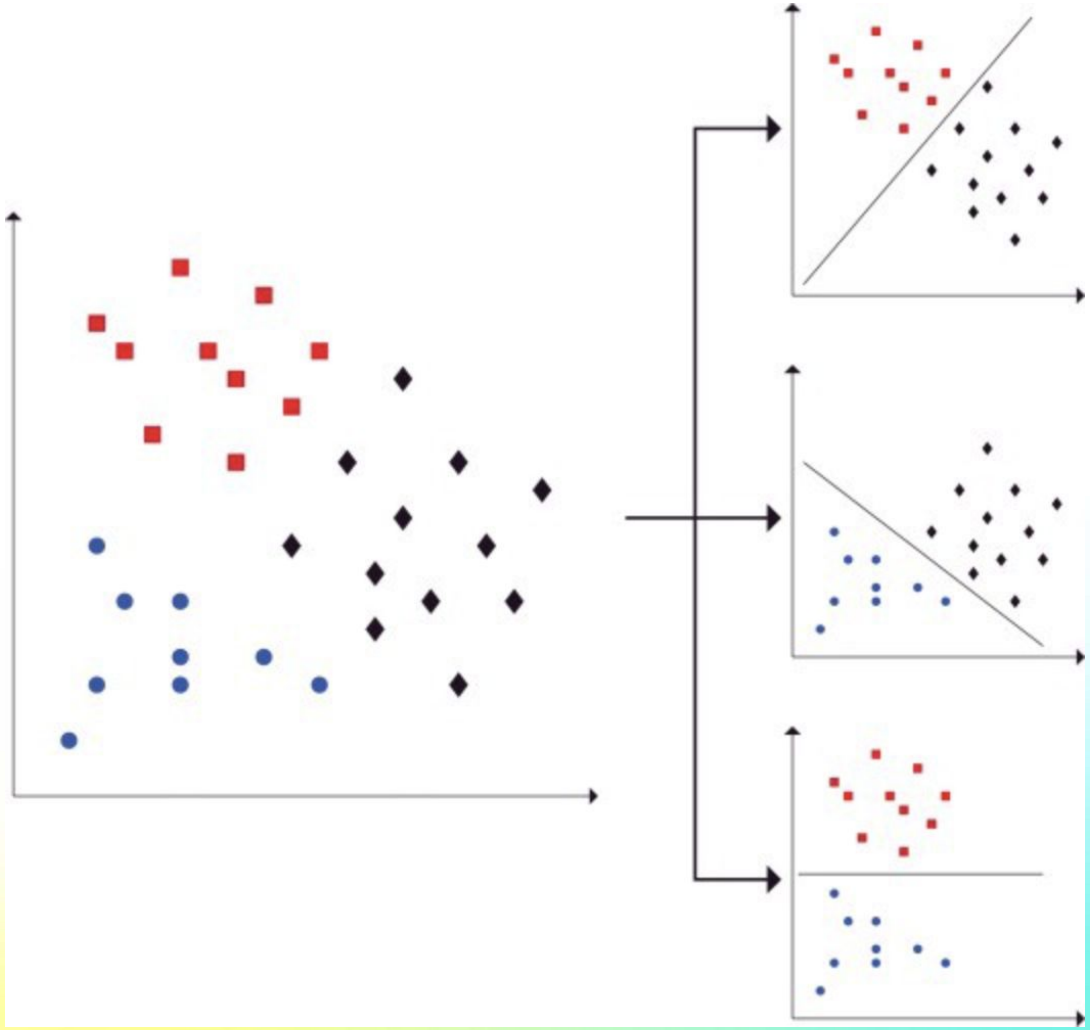
$$p_x = \frac{e^{b_k(x)}}{\sum_{j=1}^{k} e^{b_j(x)}}$$ - softmax

Многоклассовая классификация

2) Все против всех (all-vs-all, one-vs-one)

$$a(x) = \operatorname*{argmax}_{k} \sum_{i=1}^{k} \sum_{j \neq 1} [a_{ij}(x) = k]$$

a_{ij} – либо i, либо j



Метрики качества многоклассовой классификации

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

Метрики качества многоклассовой классификации

1) микро-усреднение

$\overline{TP} = \frac{1}{k}\sum_{k=1}^{k} TP_k$

$$\text{precision} = \frac{\overline{TP}}{\overline{TP} + \overline{FP}}$$

Метрики качества многоклассовой классификации

1) микро-усреднение

$$\overline{TP} = \frac{1}{k}\sum_{k=1}^{k} TP_k$$

$$\text{precision} = \frac{\overline{TP}}{\overline{TP} + \overline{FP}}$$

2) макро-усреднение

$$precision_k = \frac{TP_k}{TP_k + FP_k}$$

$$precision = \frac{1}{k}\sum_{k=1}^{k} precision_k$$

Метрики качества многоклассовой классификации

1) микро-усреднение

$$\overline{TP} = \frac{1}{k}\sum_{k=1}^{k} TP_k$$

$$\text{precision} = \frac{\overline{TP}}{\overline{TP} + \overline{FP}}$$

у маленьких классов маленький вклад

2) макро-усреднение

$$precision_k = \frac{TP_k}{TP_k + FP_k}$$

$$precision = \frac{1}{k}\sum_{k=1}^{k} precision_k$$

у всех классов равный вклад

Классификация с пересекающимися классами

multi-label classification

$y = (0,1,1,0)$

Классификация с пересекающимися классами

multi-label classification

$y = (0,1,1,0)$

1) Независимая классификация

-для каждого класса свой классификатор $a_k(x)$

Классификация с пересекающимися классами

multi-label classification

$y = (0,1,1,0)$

1) Независимая классификация

-для каждого класса свой классификатор $a_k(x)$

2) Стекинг классификаторов

-разделим выборку X на две части X_1, X_2

-на X_1 обучим независимо классификаторы для каждого класса - $b_k(x)$

-для каждого объекта из X_2 возьмем предсказания моделей $b_k(x)$ и обучим $a_k(x)$

Классификация с пересекающимися классами

Метрика качества

$y = (0,1,1,1,0,0)$

$z = (0,1,1,0,1,0)$

Классификация с пересекающимися классами

Метрика качества

$y = (0,1,1,1,0,0)$

$z = (0,1,1,0,1,0)$

hamming(y, z)

Классификация с пересекающимися классами

Метрика качества

$y = (0,1,1,1,0,0)$

$z = (0,1,1,0,1,0)$

$\text{hamming}(y, z) = 2/6 = 0.33$

Вопросы