

Тюнинг модели логистической регрессии

Елена Кантонистова

Skillbox

Перевод вероятности в класс

$$a(x) = \frac{1}{1 + e^{-(w,x)}}$$

- Если $a(x) \geq 0.5$, то класс +1
- Если $a(x) < 0.5$, то класс −1

Порог — **0.5**.

Перевод вероятности в класс

$$a(x) = \frac{1}{1 + e^{-(w,x)}}$$

- Если $a(x) \geq 0.5$, то класс +1
- Если $a(x) < 0.5$, то класс −1

Порог — 0.5. Этот порог **можно менять!**

Пример: кредитный скоринг, порог 0.5

id	Предсказанная вероятность	Правильный ответ	Предсказанный класс с порогом 0.5
1	0.6	−1	1
2	0.8	1	1
3	0.3	−1	−1
4	0.55	−1	1
5	0.1	−1	−1
6	0.96	1	1
7	0.33	1	−1
8	0.2	−1	−1
9	0.14	−1	−1
10	0.88	1	1

Матрица ошибок

Confusion Matrix

<div><div><i><u>Predict</u></i></div><div><i><u>Actual</u></i></div></div>	0	1
0	TN	FN
1	FP	TP

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

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

Точность и полнота

- Точность (или precision) — это доля верно определённых объектов среди тех, которые классификатор отнёс к положительным
- Полнота (или recall) — это доля найденных классификатором объектов положительного класса среди всех положительных объектов

Пример: кредитный скоринг, порог 0.5

Матрица ошибок:

TP = 3	FN = 1
FP = 2	TN = 4

$$precision = \frac{3}{3 + 2} = \frac{3}{5} = 0.6$$

$$recall = \frac{3}{3 + 1} = \frac{3}{4} = 0.75$$

Пример: кредитный скоринг, порог 0.25

id	Предсказанная вероятность	Правильный ответ	Предсказанный класс с порогом 0.25
1	0.6	−1	1
2	0.8	1	1
3	0.3	−1	1
4	0.55	−1	1
5	0.1	−1	−1
6	0.96	1	1
7	0.33	1	1
8	0.2	−1	−1
9	0.14	−1	−1
10	0.88	1	1

Пример: кредитный скоринг, порог 0.25

Матрица ошибок:

TP = 3	FN = 0
FP = 3	TN = 3

$$precision = \frac{4}{4 + 3} = \frac{4}{7} \approx 0.57$$

$$recall = \frac{4}{4 + 0} = \frac{4}{4} = 1$$

Метрики и порог

- При уменьшении порога полнота растёт, а точность падает
- При увеличении порога точность растёт, а полнота падает

Итоги модуля

- Узнали, что такое линейные классификаторы
- Изучили линейный классификатор под названием «логистическая регрессия»
- Узнали про градиентный спуск и реализовали его метод в python, обучили с его помощью линейную и логистическую регрессию
- Посмотрели, как применять логистическую регрессию из библиотеки sklearn и настраивать её гиперпараметры
- Узнали, что можно настраивать порог для перевода вероятности модели в классы и этим влиять на метрики