Метрики качества регрессии

Елена Кантонистова

МЕТРИКИ КАЧЕСТВА И ФУНКЦИОНАЛЫ ОШИБКИ
В ЗАДАЧАХ РЕГРЕССИИ

МЕТРИКИ КАЧЕСТВА И ФУНКЦИИ ОШИБКИ

- **Функционал (функция) ошибки –** функция, которую минимизируют в процессе обучения модели для нахождения неизвестных параметров (весов).
- **Метрика качества –** функция, которую используют для оценки качества построенной (уже обученной) модели.

МЕТРИКИ КАЧЕСТВА И ФУНКЦИИ ОШИБКИ

- **Функционал (функция) ошибки –** функция, которую минимизируют в процессе обучения модели для нахождения неизвестных параметров (весов).
- **Метрика качества –** функция, которую используют для оценки качества построенной (уже обученной) модели.

Иногда одна и та же функция может использоваться и для обучения модели (функция ошибки), и для оценки качества модели (метрика качества).

ЛИНЕЙНАЯ РЕГРЕССИЯ

Линейная регрессия:

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

Обучение линейной регрессии - минимизация среднеквадратичной ошибки:

$$\frac{1}{l}\sum_{i=1}^{l}(a(x_i) - y_i)^2 \to \min_w$$

СРЕДНЕКВАДРАТИЧНОЕ ОТКЛОНЕНИЕ: MSE (MEAN SQUARED ERROR)

Среднеквадратичное отклонение:

$$MSE(a, X) = \frac{1}{l}\sum_{i=1}^{l}\boldsymbol{(a(x_i) - y_i)^2}$$

СРЕДНЕКВАДРАТИЧНОЕ ОТКЛОНЕНИЕ: MSE (MEAN SQUARED ERROR)

Среднеквадратичное отклонение:

$$MSE(a, X) = \frac{1}{l}\sum_{i=1}^{l}\boldsymbol{(a(x_i) - y_i)^2}$$

Плюсы:

- Позволяет сравнивать модели
- Подходит для контроля качества во время обучения

СРЕДНЕКВАДРАТИЧНОЕ ОТКЛОНЕНИЕ: MSE

Среднеквадратичное отклонение:

$$MSE(a, X) = \frac{1}{l}\sum_{i=1}^{l}\boldsymbol{(a(x_i) - y_i)^2}$$

Плюсы:

- Позволяет сравнивать модели
- Подходит для контроля качества во время обучения

Минусы:

- Плохо интерпретируется, т.к. не сохраняет единицы измерения (если целевая переменная – кг, то MSE измеряется в кг в квадрате)
- Тяжело понять, насколько хорошо данная модель решает задачу, так как MSE не ограничена сверху.

RMSE (ROOT MEAN SQUARED ERROR)

Корень из среднеквадратичной ошибки:

$$RMSE(a, X) = \sqrt{\frac{1}{l}\sum_{i=1}^{l}(a(x_i) - y_i)^2}$$

Плюсы:

- Все плюсы MSE
- **Сохраняет единицы измерения (в отличие от MSE)**

Минусы:

- Тяжело понять, насколько хорошо данная модель решает задачу, так как RMSE не ограничена сверху.

КОЭФФИЦИЕНТ ДЕТЕРМИНАЦИИ (R^2)

Коэффициент детерминации:

$$R^2(a, X) = 1 - \frac{\sum_{i=1}^{l} (a(x_i) - y_i)^2}{\boldsymbol{\sum_{i=1}^{l} (y_i - \overline{y})^2}},$$

где $\overline{y} = \frac{1}{l} \sum_{i=1}^{l} y_i$.

КОЭФФИЦИЕНТ ДЕТЕРМИНАЦИИ (R^2)

Коэффициент детерминации:

$$R^2(a, X) = 1 - \frac{\sum_{i=1}^{l}(a(x_i) - y_i)^2}{\sum_{i=1}^{l}(y_i - \overline{y})^2},$$

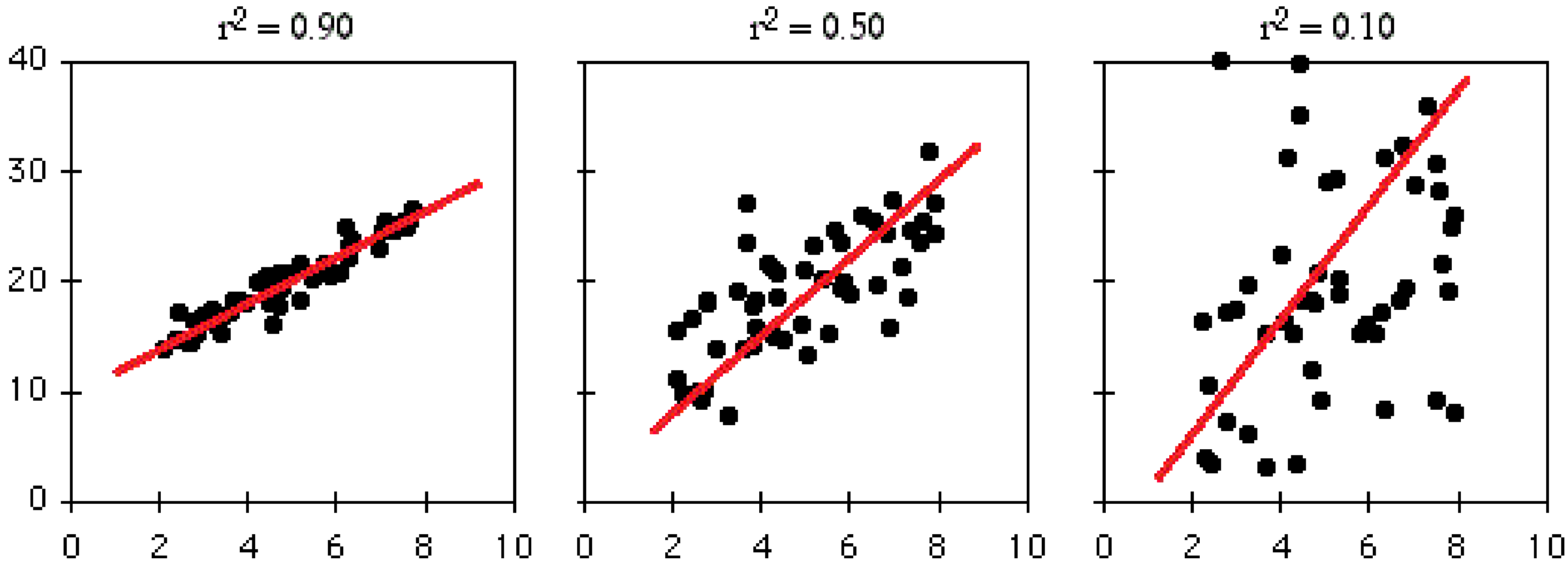
где $\overline{y} = \frac{1}{l}\sum_{i=1}^{l} y_i$.

Коэффициент детерминации ***это доля дисперсии целевой переменной, объясняемая моделью***.

- Чем ближе R^2 к 1, тем лучше модель объясняет данные
- Чем ближе R^2 к 0, тем ближе модель к константному предсказанию
- Отрицательный R^2 говорит о том, что модель плохо решает задачу

КОЭФФИЦИЕНТ ДЕТЕРМИНАЦИИ (R^2)

$$\boldsymbol{R^2 \leq 1}$$



MAE (MEAN ABSOLUTE ERROR)

Средняя абсолютная ошибка:

$$MAE(a, X) = \frac{1}{l}\sum_{i=1}^{l} \boldsymbol{|a(x_i) - y_i|}$$

MAE (MEAN ABSOLUTE ERROR)

Средняя абсолютная ошибка:

$$MAE(a, X) = \frac{1}{l}\sum_{i=1}^{l} |\boldsymbol{a}(\boldsymbol{x}_{\boldsymbol{i}}) - \boldsymbol{y}_{\boldsymbol{i}}|$$

Плюсы:

- Менее чувствителен к выбросам, чем MSE

MAE (MEAN ABSOLUTE ERROR)

Средняя абсолютная ошибка:

$$MAE(a, X) = \frac{1}{l}\sum_{i=1}^{l} |\boldsymbol{a(x_i) - y_i}|$$

Плюсы:

- Менее чувствителен к выбросам, чем MSE

Минусы:

- MAE - не дифференцируемый функционал

MSLE (MEAN SQUARED LOGARITHMIC ERROR)

Среднеквадратичная логарифмическая ошибка:

$$\mathrm{M}SLE(a, X) = \frac{1}{l}\sum_{i=1}^{l}\left(\mathbf{log(a(x_i) + 1) - log(y + 1)}\right)^{\mathbf{2}}$$

- Подходит для задач с неотрицательной целевой переменной ($y \geq 0$)
- Штрафует за отклонения в порядке величин
- Штрафует заниженные прогнозы сильнее, чем завышенные

MAPE

MAPE – Mean Absolute Percentage Error:

$$MAPE(a, X) = \frac{1}{l}\sum_{i=1}^{l}\frac{|\boldsymbol{y_i} - \boldsymbol{a}(\boldsymbol{x_i})|}{|\boldsymbol{y_i}|}$$

MAPE измеряет относительную ошибку.

MAPE

$$MAPE(a,X) = \frac{1}{l}\sum_{i=1}^{l}\frac{|\boldsymbol{y_i} - \boldsymbol{a}(\boldsymbol{x_i})|}{|\boldsymbol{y_i}|}$$

Плюсы:

- $MAPE \geq 0$
- Хорошо интерпретируема: например, MAPE=0.16 означает, что ошибка модели в среднем составляет 16% от фактических значений.

MAPE

$$MAPE(a, X) = \frac{1}{l}\sum_{i=1}^{l}\frac{|\boldsymbol{y_i} - \boldsymbol{a}(\boldsymbol{x_i})|}{|\boldsymbol{y_i}|}$$

Плюсы:

- $MAPE \geq 0$
- Хорошо интерпретируема: например, MAPE=0.16 означает, что ошибка модели в среднем составляет 16% от фактических значений.

Минусы:

- По-разному относится к недо- и перепрогнозу. Например, если правильный ответ $y = 10$, а прогноз $a(x) = 20$, то ошибка $\frac{|10-20|}{|10|} = \mathbf{1}$, а если ответ $y = 30$, то ошибка $\frac{|30-20|}{|30|} = \frac{1}{3} \approx \mathbf{0.33}$.

SMAPE

SMAPE – *Symmetric Mean Absolute Percentage Error* (симметричный вариант MAPE):

$$SMAPE(a, X) = \frac{1}{l} \sum_{i=1}^{l} \frac{|\boldsymbol{y_i} - \boldsymbol{a}(\boldsymbol{x_i})|}{(|\boldsymbol{y_i}| + |\boldsymbol{a}(\boldsymbol{x_i})|)/\mathbf{2}}$$

SMAPE – попытка сделать симметричным прогноз (то есть дать одинаковую ошибку для недо- и перепрогноза).

SMAPE

SMAPE – *Symmetric Mean Absolute Percentage Error* (симметричный вариант MAPE):

$$SMAPE(a, X) = \frac{1}{l}\sum_{i=1}^{l} \frac{|\boldsymbol{y_i} - \boldsymbol{a}(\boldsymbol{x_i})|}{(|\boldsymbol{y_i}| + |\boldsymbol{a}(\boldsymbol{x_i})|)/\boldsymbol{2}}$$

SMAPE – попытка сделать симметричным прогноз (то есть дать одинаковую ошибку для недо- и перепрогноза).

<u>Проверим:</u>

Пусть правильный ответ $\boldsymbol{y = 10}$, а прогноз $\boldsymbol{a(x) = 20}$, то ошибка $\frac{|10-20|}{|10+20|/2} = \frac{2}{3} \approx \boldsymbol{0.67}$, а если ответ $\boldsymbol{y = 30}$, то ошибка $\frac{|30-20|}{|30+20|/2} = \frac{2}{5} = \boldsymbol{0.4}$.

SMAPE

SMAPE – попытка сделать симметричным прогноз (то есть дать одинаковую ошибку для недо- и перепрогноза).

<u>Проверим:</u>

Пусть правильный ответ $\boldsymbol{y = 10}$, а прогноз $\boldsymbol{a(x) = 20}$, то ошибка $\frac{|10-20|}{|10+20|/2} = \frac{2}{3} \approx \boldsymbol{0.67}$, а если ответ $\boldsymbol{y = 30}$, то ошибка $\frac{|30-20|}{|30+20|/2} = \frac{2}{5} = \boldsymbol{0.4}$.

Ошибки стали меньше отличаться друг от друга, но всё-таки не равны.

SMAPE

SMAPE – попытка сделать симметричным прогноз (то есть дать одинаковую ошибку для недо- и перепрогноза).

“Сейчас уже в среде прогнозистов сложилось более-менее устойчивое понимание, что SMAPE не является хорошей ошибкой. Тут дело не только в завышении прогнозов, но ещё и в том, что наличие прогноза в знаменателе позволяет манипулировать результатами оценки.” (см. источник)