Анализ категориальных данных

Лекция 2. Модели бинарного выбора: оценка качества модели

2 февраля 2021

Вопрос

Что из себя представляет confusion matrix? Как ее построить?

Вопрос

Что из себя представляет confusion matrix? Как ее построить?

Ответ

1. Сначала нужно сохранить предсказанные моделью вероятности $P(Y = 1)$
2. Далее выбрать порог отсечения: к примеру, если $P(Y = 1)$ более 0.5 отнести наблюдение к классу 1, и в противном случае – к классу 0.
3. Далее на основе предсказанных и наблюдаемых значений можно построить аналог таблицы сопряженности

Вопрос

Рассмотрим элементы confusion matrix подробнее.

Вопрос

Рассмотрим элементы confusion matrix подробнее.

Ответ

$$\begin{pmatrix} & data: Y=1 & Y=0 \\ prediction: Y=1 & TP & FP \\ prediction: Y=0 & FN & TN \end{pmatrix}$$, где

TP – истинно «положительные» значения (в реальности относится к классу 1 и классифицировано моделью так же)

TN – по аналогии: истинно «отрицательные» значения

FP – допущена ошибка классификатором: отнесли к классу 1 («положительные»), а на самом деле – класс 0

FN – допущена ошибка классификатором: отнесли к классу 0 («отрицательные»), а на самом деле – класс 1

Вопрос

Определите по этой confusion matrix ошибку I рода, ошибку II рода и мощность критерия.

Вопрос

Определите по этой confusion matrix ошибку I рода, ошибку II рода и мощность критерия.

Ответ

$$\begin{pmatrix} & data: Y=1 & Y=0 \\ prediction: Y=1 & TP & FP \\ prediction: Y=0 & FN & TN \end{pmatrix},$$

$$\text{ошибка I рода} = P(reject|H0) = \frac{FP}{FP+TN}$$

$$\text{ошибка II рода} = P(NOTreject|H1) = \frac{FN}{FN+TP}$$

$$\text{мощность} = P(reject|H1) = \frac{TP}{FN+TP}$$

Чтобы confusion matrix не смогла Вас confuse:

1. Когда считаете ошибку I рода, вспоминайте, что теперь массив сужается только до H0 (класс 0 по ИСХОДНЫМ ДАННЫМ): отвержение при условии верной H0. Мысленно оставляйте в матрице только тот столбец, который соответствует (data: Y = 1), то есть, TN + FP. А дальше зададимся вопросом, когда допускается ошибка? FP – это, конечно, же ошибка, поэтому и получаем $\frac{FP}{FP+TN}$
2. По аналогии делайте и при расчете ошибки II рода: только теперь Вас интересует подмассив «класс 1»

Вопрос

Что такое меры чувствительности (sensitivity) и специфичности (specificity)?

Вопрос

Что такое меры чувствительности (sensitivity) и специфичности (specificity)?

Ответ

Когда считаем эти меры, нас всегда будет интересовать, какую долю наблюдений мы классифицировали моделью ВЕРНО (относительно исходных данных). Осталось только запомнить, что чувствительность – это про верные «положительные» наблюдения, а специфичность – про верные «отрицательные».

$$\text{Sensitivity} = \frac{TP}{TP + FN}; \text{Specificity} = \frac{TN}{TN + FP}$$

Несложно заметить, что

Sensitivity – это мощность критерия (которую мы всегда хотим максимизировать).
Specificity – это (1 – ошибка I рода), эту величину тоже хочется максимизировать.
Однако одновременно это сделать на практике сложно, для того, чтобы найти подходящее пороговое значение (насколько это возможно, максимизирующее мощность и миниминизирующее ошибку I рода) нам пригодится ROC. См. полезный интерактив с бегунком по ROC – здесь.

Меры качества модели: R^2

Для логистических моделей, так же как и для классических линейных, существуют R^2, только они псевдо-R^2. Они основаны на функции правдоподобия модели и НЕ могут интерпретироваться как доля объясненной вариации. Подробнее про разные варианты pseudo-R^2 можно посмотреть здесь.