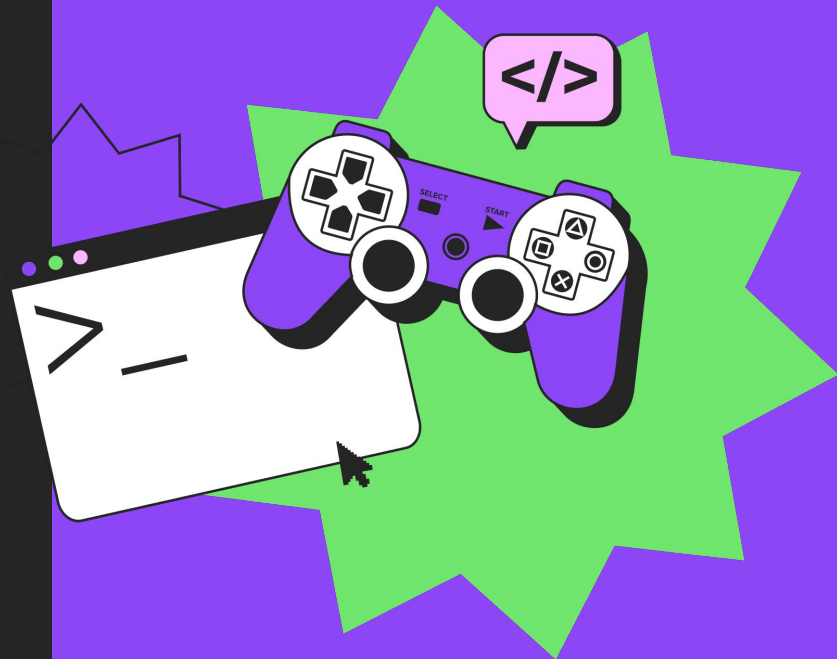


Оценка моделей





Александр Немальцев






Руководитель направления, СБЕР

Более 5 лет работаю в сфере анализа данных и Data Science

- ✨ Внедрил автоматизированную систему анализа отзывов клиентов в целях определения проблемных зон во взаимодействии с клиентом и принятия на её основе управленческих решений
- ✨ Разрабатываю модели рекомендаций
- ✨ Работал в таких крупных компаниях, как SAS, X5 Group



Что будет на уроке сегодня

-  Метрики регрессии
-  Метрики для оценки прогноза временных рядов
-  Метрики бинарной классификации
-  Метрики многоклассовой классификации
-  Метрики кластеризации





Основные виды моделей

Обучение с учителем

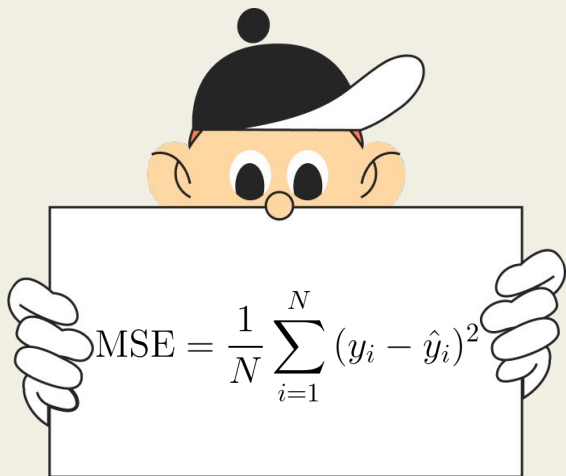
- задача восстановления регрессии
- задача классификации

Обучение без учителя



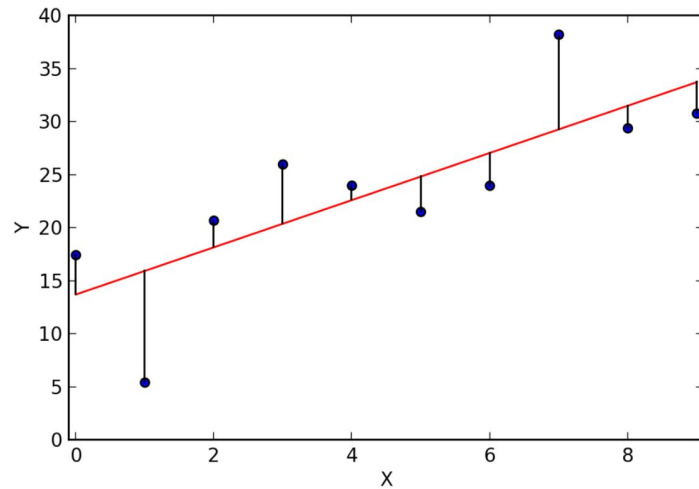


Обучение с учителем. Задача регрессии



A cartoon character with an orange face, large eyes, and a black and white cap is holding a white rectangular sign. The sign displays the Mean Squared Error (MSE) formula.

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$





Обучение с учителем.

Задача регрессии

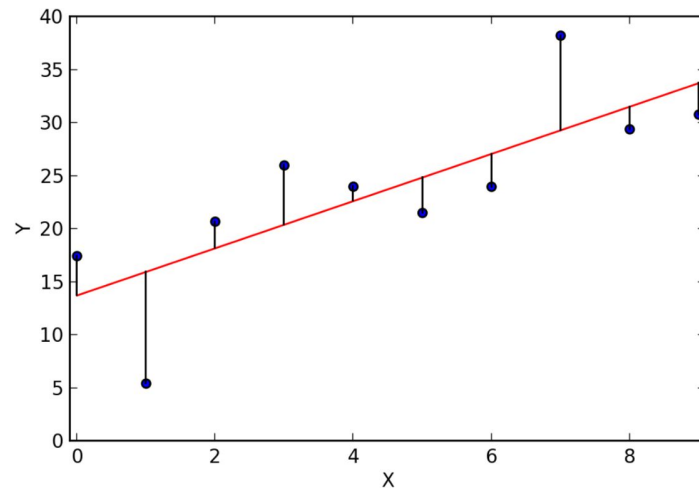
Плюсы:

- ♥ Прост в реализации
- ♥ Позволяет оценить качество модели относительно среднего предсказания

Минусы:

- ♥ Плохо интерпретировать

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$





Обучение с учителем. Задача регрессии

Плюсы:

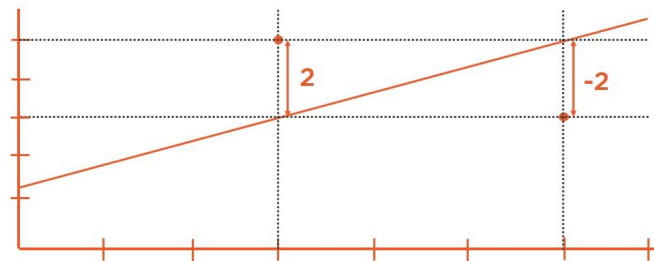
- ♥ Прост в реализации
- ♥ Численно понимаем, как ошибается модель

Минусы:

- ♥ Не ясно, в какую сторону модель ошибается больше
- ♥ Диапазон ответов не нормирован

$$RMSE = \sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2}$$

Root Mean Squared Error (RMSE)



Mean Squared Error (MSE):

$$\frac{2^2 + (-2)^2}{2} = \frac{4 + 4}{2} = 4$$



Обучение с учителем.

Задача регрессии

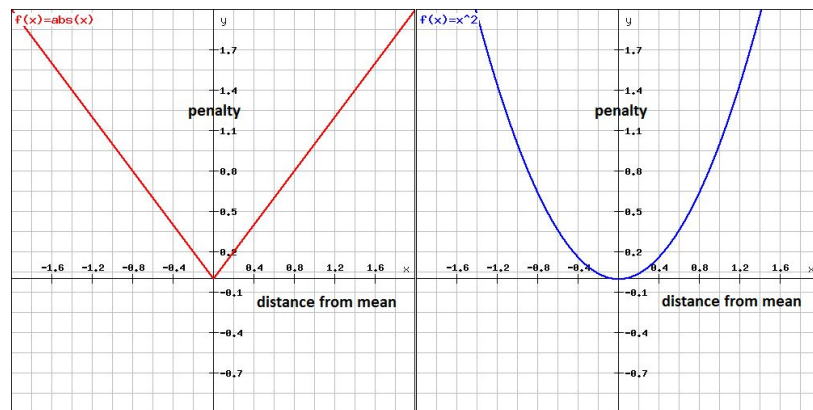
Плюсы:

- ♥ Не так чувствителен к выбросам
- ♥ Штрафует более высокие отклонения

Минусы:

- ♥ Диапазон ответов не нормирован
- ♥ Производная в точке 0 не определена

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$





Обучение с учителем. Задача регрессии

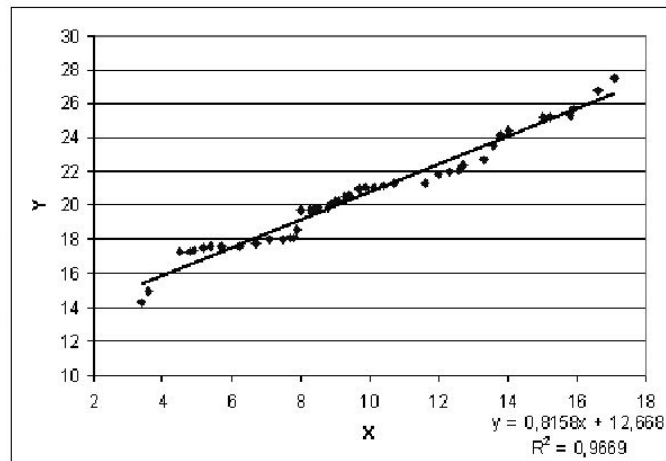
Плюсы:

- ♥ Хорошо объясняет полученную модель
- ♥ Принимает ограниченный диапазон

Минусы:

- ♥ Не всегда модель с высоким показателем R^2 работает корректно

$$R^2 \equiv 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2},$$





Обучение с учителем. Задача регрессии

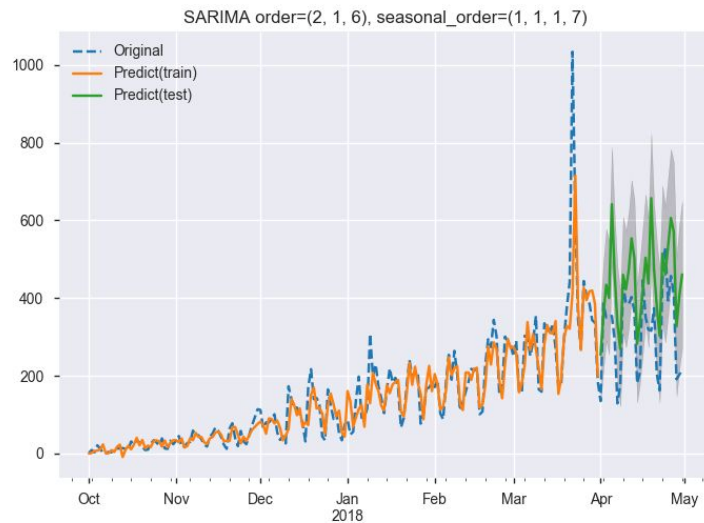
Плюсы:

♥ Понимание, в какую сторону смещён прогноз

Минусы:

♥ Не учитывает ошибку на каждом шаге прогноза

$$BIAS = \frac{Actual - Forecast}{ABS(Actual - Forecast)}$$





Обучение с учителем. Задача регрессии

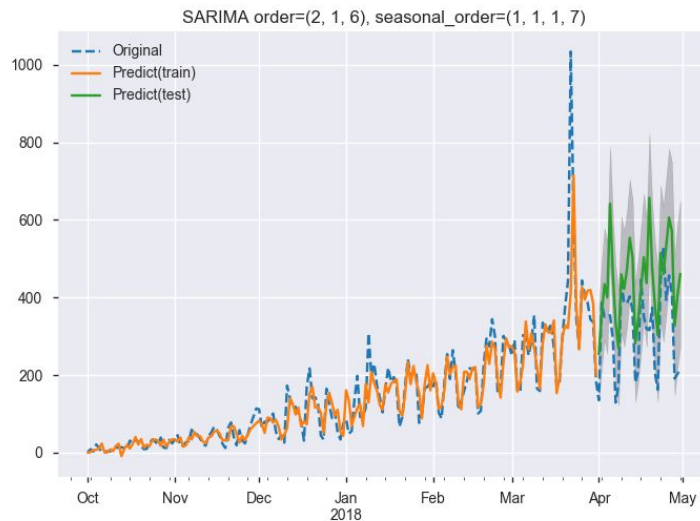
Плюсы:

♥ Учитывает ошибку прогноза на каждом шаге

Минусы:

- ♥ Нет понимания, как смещён прогноз
- ♥ Не учтены суммарные продажи за весь период прогноза

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|$$





Обучение с учителем. Задача регрессии

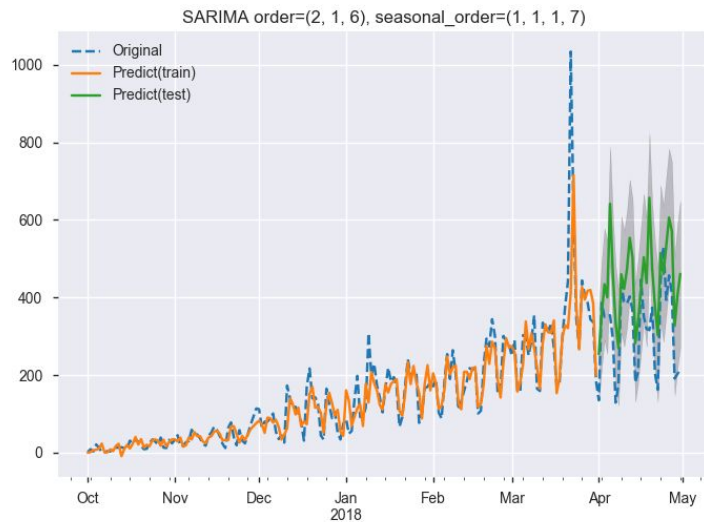
Плюсы:

- ♥ Учитывает ошибку прогноза на каждом шаге
- ♥ Учитывает сумму продаж за весь период

Минусы:

- ♥ Нет понимания, как смещён прогноз

$$W A P E = \frac{\sum_{t=1}^n |A_t - F_t|}{\sum_{t=1}^n |A_t|}$$





Обучение с учителем. Задача классификации

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

		True Class	
		Positive	Negative
Predicted Class	Positive	TP	FP
	Negative	FN	TN

Плюсы:

- ♥ Прост в реализации
- ♥ Показывает процент точных попаданий

Минусы:

- ♥ Не применим в задачах с сильным дисбалансом классов




Обучение с учителем. Задача классификации


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

		True Class	
		Positive	Negative
Predicted Class	Positive	TP	FP
	Negative	FN	TN

Плюсы:

 Устойчив к дисбалансу классов

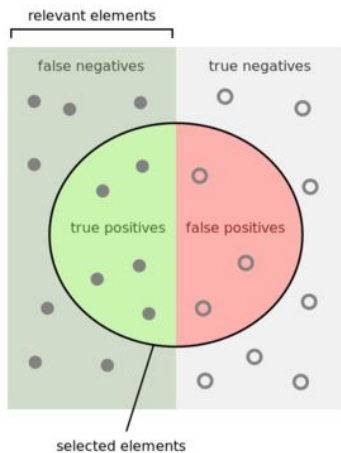
Минусы:

 Смотрит в основном на положительный класс



Обучение с учителем. Задача классификации

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



How many selected
items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant
items are selected?

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

Плюсы:

♥ Устойчив к дисбалансу классов

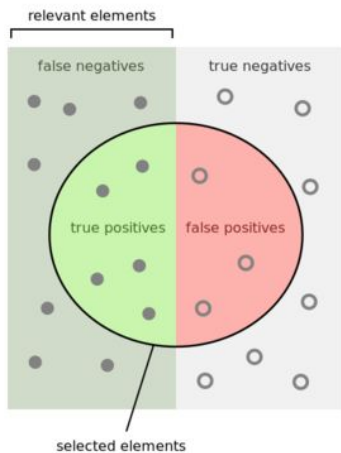
Минусы:

♥ Смотрит в основном на положительный класс



Обучение с учителем. Задача классификации

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



How many selected
items are relevant?

Precision =



How many relevant
items are selected?

Recall =



Плюсы:



Объединяет 2 метрики в одну



Позволяет задавать веса для одной из метрик

Минусы:



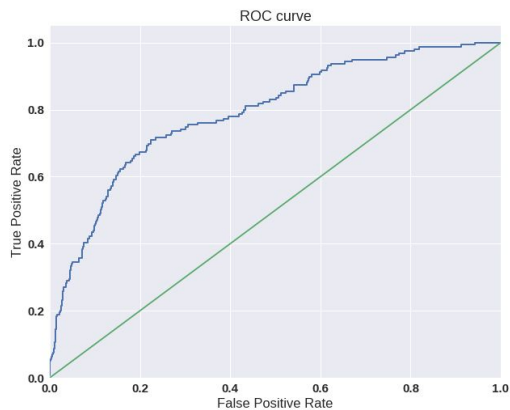
Не работает с вероятностными моделями



Обучение с учителем. Задача классификации

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

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



Плюсы:

- ♥ Ранжирующая метрика
- ♥ Устойчив к дисбалансу классов
- ♥ Работает с вероятностными моделями

Минусы:

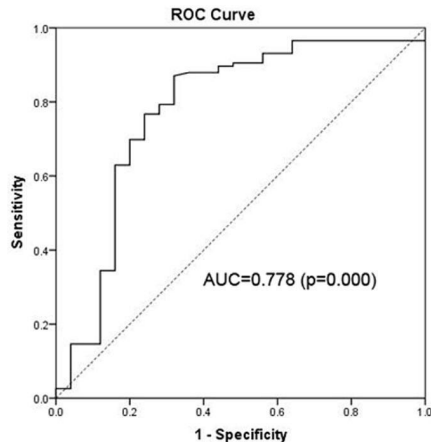
- ♥ Не во всех задачах применима



Обучение с учителем. Задача классификации

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

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



IgG (mg/ml)	Sensitivity	Specificity
6.7350	.888	.560
6.8450	.879	.560
6.9100	.879	.600
6.9350	.879	.680
6.9550	.871	.680
6.9900	.862	.680
7.0550	.853	.680
7.1400	.845	.680
7.2100	.836	.680

Плюсы:

- ♥ Ранжирующая метрика
- ♥ Устойчив к дисбалансу классов
- ♥ Работает с вероятностными моделями

Минусы:

- ♥ Не во всех задачах применима



Обучение с учителем. Задача классификации

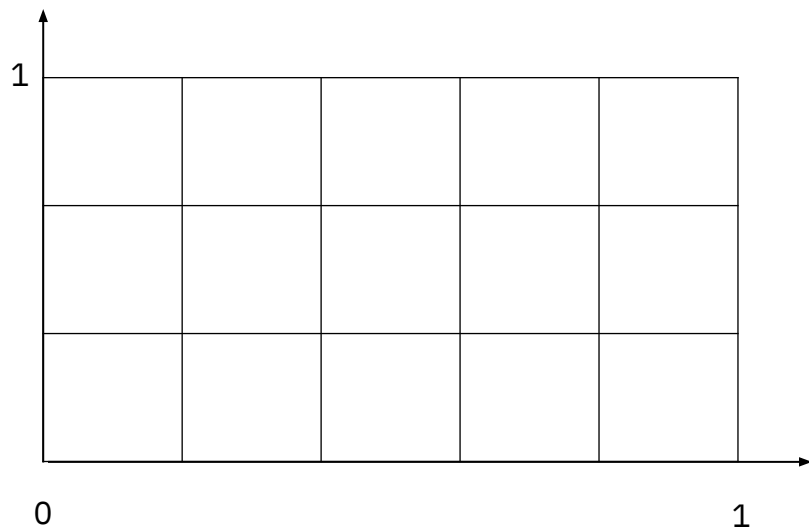
y	prediction	predict_proba
0	1	0.6
1	1	0.9
0	0	0.05
0	1	0.99
1	0	0.2
1	1	0.6
0	0	0.3
0	1	0.51



y	prediction	predict_proba
0	1	0.99
1	1	0.9
0	1	0.6
1	1	0.6
0	1	0.51
0	0	0.3
1	0	0.2
0	0	0.05



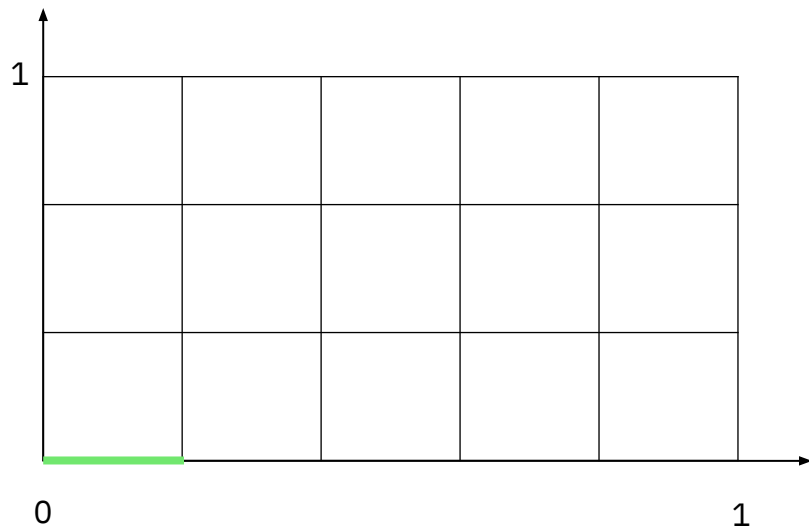
Обучение с учителем. Задача классификации



y	prediction	predict_proba
0	1	0.99
1	1	0.9
0	1	0.6
1	1	0.6
0	1	0.51
0	0	0.3
1	0	0.2
0	0	0.05



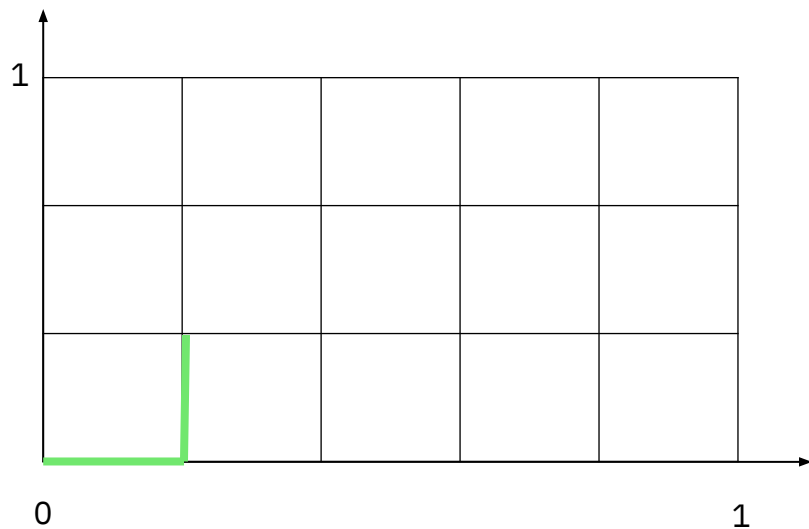
Обучение с учителем. Задача классификации



y	prediction	predict_proba
0	1	0.99
1	1	0.9
0	1	0.6
1	1	0.6
0	1	0.51
0	0	0.3
1	0	0.2
0	0	0.05



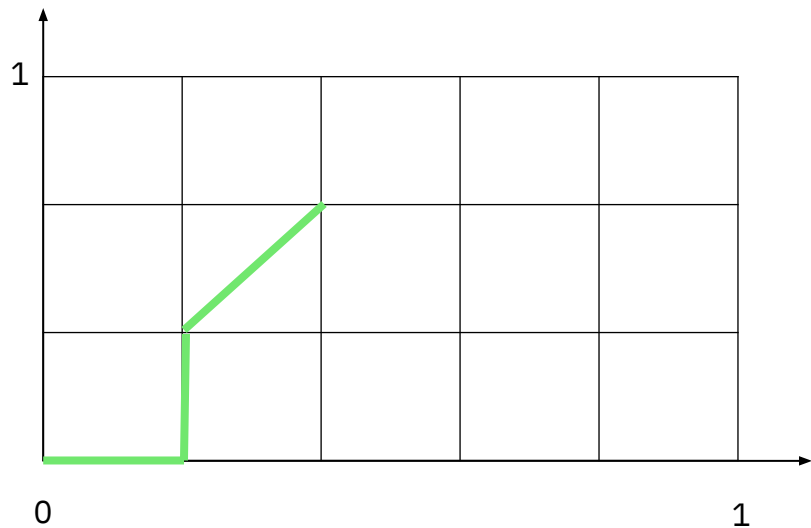
Обучение с учителем. Задача классификации



y	prediction	predict_proba
0	1	0.99
1	1	0.9
0	1	0.6
1	1	0.6
0	1	0.51
0	0	0.3
1	0	0.2
0	0	0.05



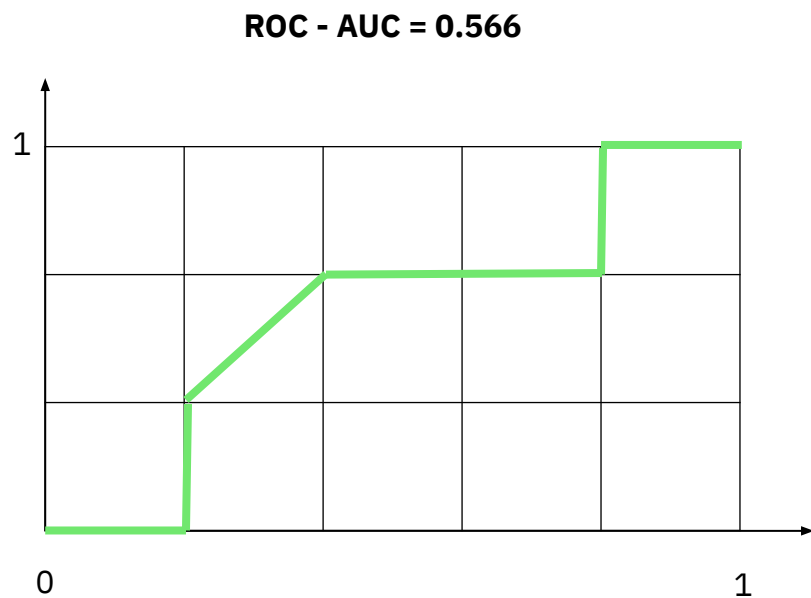
Обучение с учителем. Задача классификации



y	prediction	predict_proba
0	1	0.99
1	1	0.9
0	1	0.6
1	1	0.6
0	1	0.51
0	0	0.3
1	0	0.2
0	0	0.05



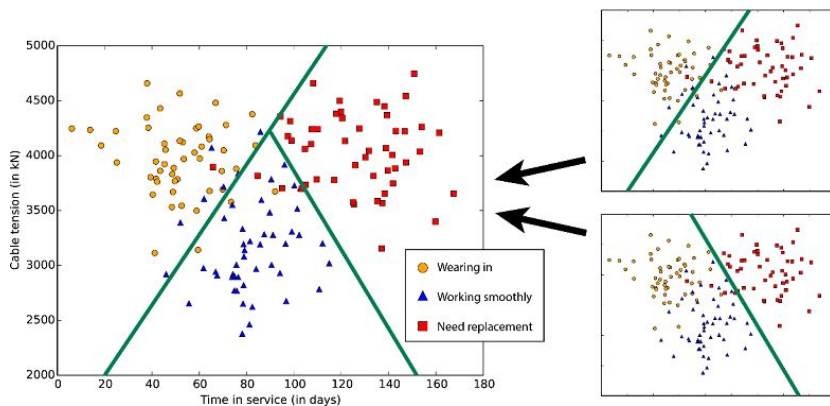
Обучение с учителем. Задача классификации



y	prediction	predict_proba
0	1	0.99
1	1	0.9
0	1	0.6
1	1	0.6
0	1	0.51
0	0	0.3
1	0	0.2
0	0	0.05



Обучение с учителем. Задача классификации



Многоклассовая классификация

1. Усредняем элементы матрицы ошибок (TP, FP, TN, FN) между бинарными классификаторами. Затем по одной усреднённой матрице ошибок считаем Precision, Recall, F-меру. Это называют **микроусреднением**.
2. Считаем Precision, Recall для каждого классификатора отдельно, а потом усредняем. Это называют **макроусреднением**.



Обучение с учителем. Задача классификации

Коэффициент силуэта:

$$S(x_i) = \frac{B(x_i) - A(x_i)}{\max(B(x_i), A(x_i))}$$

Гомогенность:

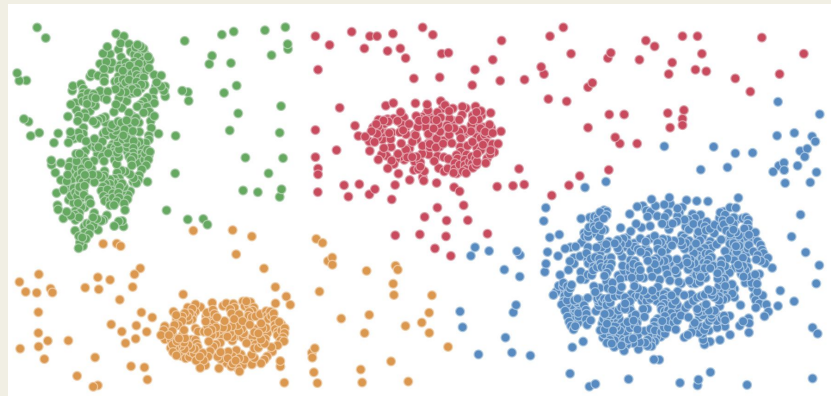
$$Homogeneity = 1 - \frac{H_{class|clust}}{H_{class}}$$

Полнота:

$$Completeness = 1 - \frac{H_{clust|class}}{H_{clust}}$$

V-мера:

$$V_\beta = \frac{(1 + \beta) \cdot Homogeneity \cdot Completeness}{\beta \cdot Homogeneity + Completeness}$$





Что мы сегодня узнали и чему научились



Метрики регрессии



Метрики для оценки прогноза временных рядов



Метрики бинарной классификации



Метрики многоклассовой классификации



Метрики кластеризации



Практическое задание

Предположим, что мы решаем задачу определения мошеннических операций. Мы разделили выборку, обучили модель и теперь хотим её оценить.

Требуется рассчитать следующие ключевые метрики:

1. Recall
2. Precision
3. Accuracy
4. f1-score

Как изменятся метрики, если в качестве бейзлайна (самой простой модели) возьмём константную модель, которая всегда выдаёт ответ 1?

y	prediction	predict_proba
0	1	0.7
0	0	0.1
1	1	0.9
1	0	0.5
0	0	0.4



Вопросы?

Вопросы?



Вопросы?

