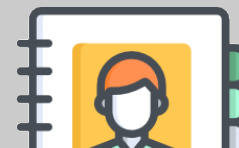


# Метрики оценки качества



## План

1. Регрессия

2. Классификация

3. Пример

# 1. Регрессия



# Метрики качества

RMSE

MAE

MAPE

SMAPE

logloss

# ROOT MEAN SQUARED ERROR

- Корень из среднего квадратичного отклонения прогноза от исходного значения
- Сильно штрафует за бОльшие по модулю отклонения

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

# MEAN AVERAGE ERROR

- Отклонение прогноза от исходного значения
- Усредненное по всем наблюдениям

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

# MEAN AVERAGE PERCENTAGE ERROR

- Ошибка прогнозирования оценивается в процентах

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

# SYMMETRIC MEAN AVERAGE PERCENTAGE ERROR

- Ошибка прогнозирования оценивается в процентах

$$\text{SMAPE} = \frac{100\%}{n} \sum_{t=1}^n \frac{|F_t - A_t|}{(|A_t| + |F_t|)/2}$$



# SYMMETRIC MEAN AVERAGE PERCENTAGE ERROR

- Ошибка прогнозирования оценивается в процентах

$$\text{SMAPE} = \frac{100\%}{n} \sum_{t=1}^n \frac{|F_t - A_t|}{(|A_t| + |F_t|)/2}$$

$$\text{SMAPE} = \frac{100\%}{n} \sum_{t=1}^n \frac{|F_t - A_t|}{|A_t| + |F_t|}$$

# SYMMETRIC MEAN AVERAGE PERCENTAGE ERROR

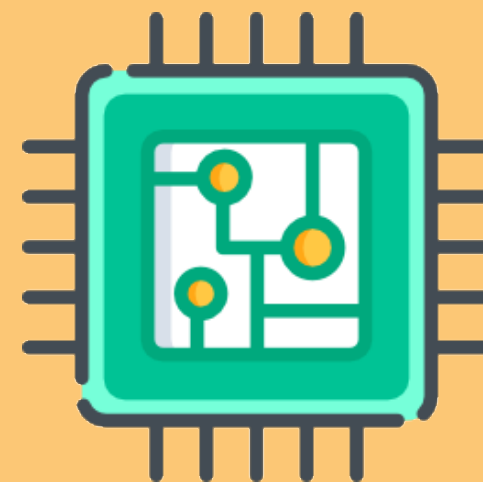
- По-разному штрафует за перепрогнозирование и недопрогнозирование
- Перепрогнозирование:  
 $A_t = 100, F_t = 110 \sim \text{SMAPE} = 4.76\%$
- Недопрогнозирование:  
 $A_t = 100, F_t = 90 \sim \text{SMAPE} = 5.26\%$

# LogLoss

- Логарифмическая ошибка
- Хорошо оценивает вероятность

$$LogLoss = -\frac{1}{n} \sum_{i=0}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

## 2. Классификации



# Метрики

Accuracy

Precision

Recall

F-measure

ROC-AUC

# Accuracy

Доля правильных ответов при классификации

# Accuracy

Доля правильных ответов при классификации

target:     1 0 1 0 0 0 0 1 0 0

# Accuracy

Доля правильных ответов при классификации

target: 1 0 1 0 0 0 0 1 0 0

predicted: 0 0 1 0 0 0 0 1 1 0



# Accuracy

Доля правильных ответов при классификации

target: 1 0 1 0 0 0 0 1 0 0

predicted: 0 0 1 0 0 0 0 1 1 0

# Accuracy

Доля правильных ответов при классификации

target: 1 0 1 0 0 0 0 1 0 0

predicted: 0 0 1 0 0 0 0 1 1 0

accuracy = 8/10 = 0.8

# Precision & Recall

- Precision – точность
- Recall – полнота

# Сбитые самолеты



# Сбитые самолеты



$y = (0000101101)$

$\hat{y} = (0110100101)$  ←



# Precision

Precision – точность выстрелов

Количество сбитых самолётов

Количество выстрелов

$y = (0\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 1)$

$\hat{y} = (0\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 1)$



# Recall

Recall – «полнота» сбивания

Количество сбитых самолётов

Общее количество самолетов

$$y = (0\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 1)$$

$$\hat{y} = (0\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 1)$$



# Confusion matrix

		True Class	
		Positive (P)	Negative (N)
Predicted Class	Positive (+)	True Positive Count (TP)	False Positive Count (FP)
	Negative (-)	False Negative Count (FN)	True Negative Count (TN)

$$\text{Precision} = \frac{tp}{tp + fp}$$

$$\text{Recall} = \frac{tp}{tp + fn}$$



## F-measure (F-score, F1)

- Среднее гармоническое между precision и recall
- Значение F-measure ближе к меньшему из precision, recall

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

# ROC-AUC

- Применяется для оценки «вероятностной» классификации \*
- «Качество» ранжирования объектов по вероятности принадлежности к целевому классу
- Доля «правильно» отранжированных пар
- Вероятность встретить объект целевого класса раньше, чем объект нецелевого класса

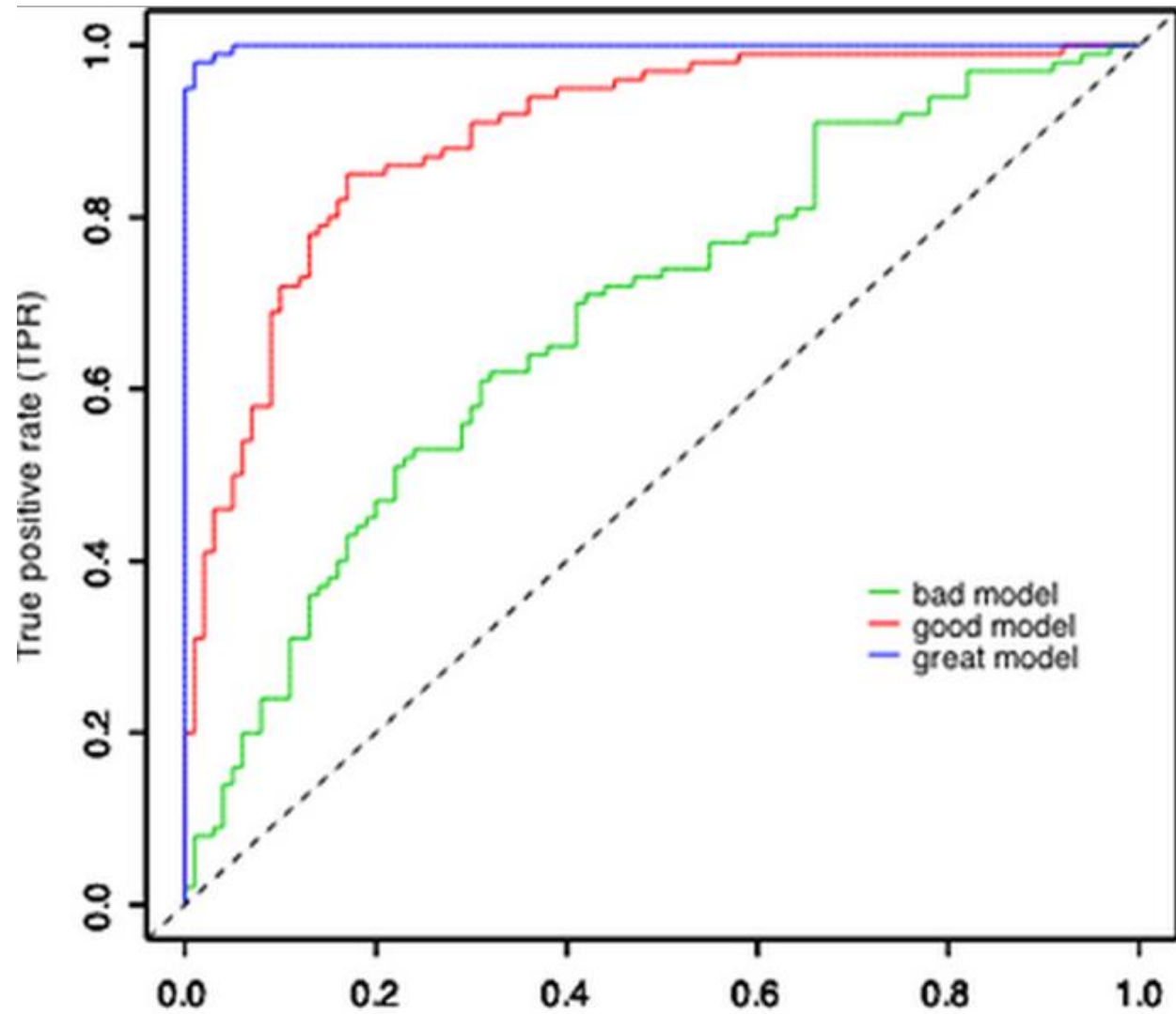
# ROC

		True Class	
		Positive (P)	Negative (N)
Predicted Class	Positive (+)	True Positive Count (TP)	False Positive Count (FP)
	Negative (-)	False Negative Count (FN)	True Negative Count (TN)

$$TPR = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

$$FPR = \frac{\text{False positives}}{\text{False positives} + \text{True negatives}}$$

# ROC



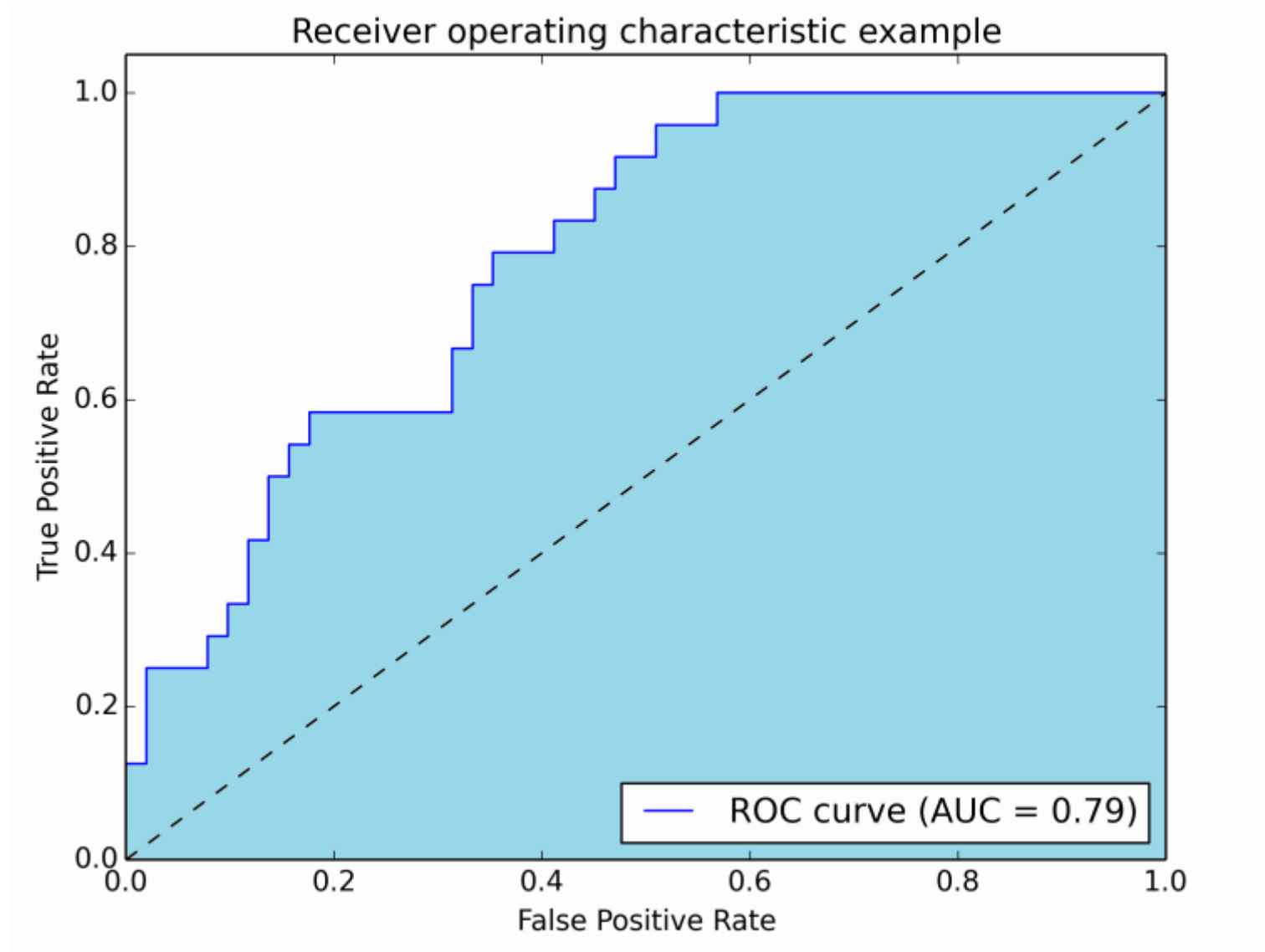
# ROC-AUC

- Как оценить кривую численно?

# ROC-AUC

- Как оценить кривую численно?
- Измерить площадь под кривой – Area Under Curve

# ROC-AUC



# ROC-AUC

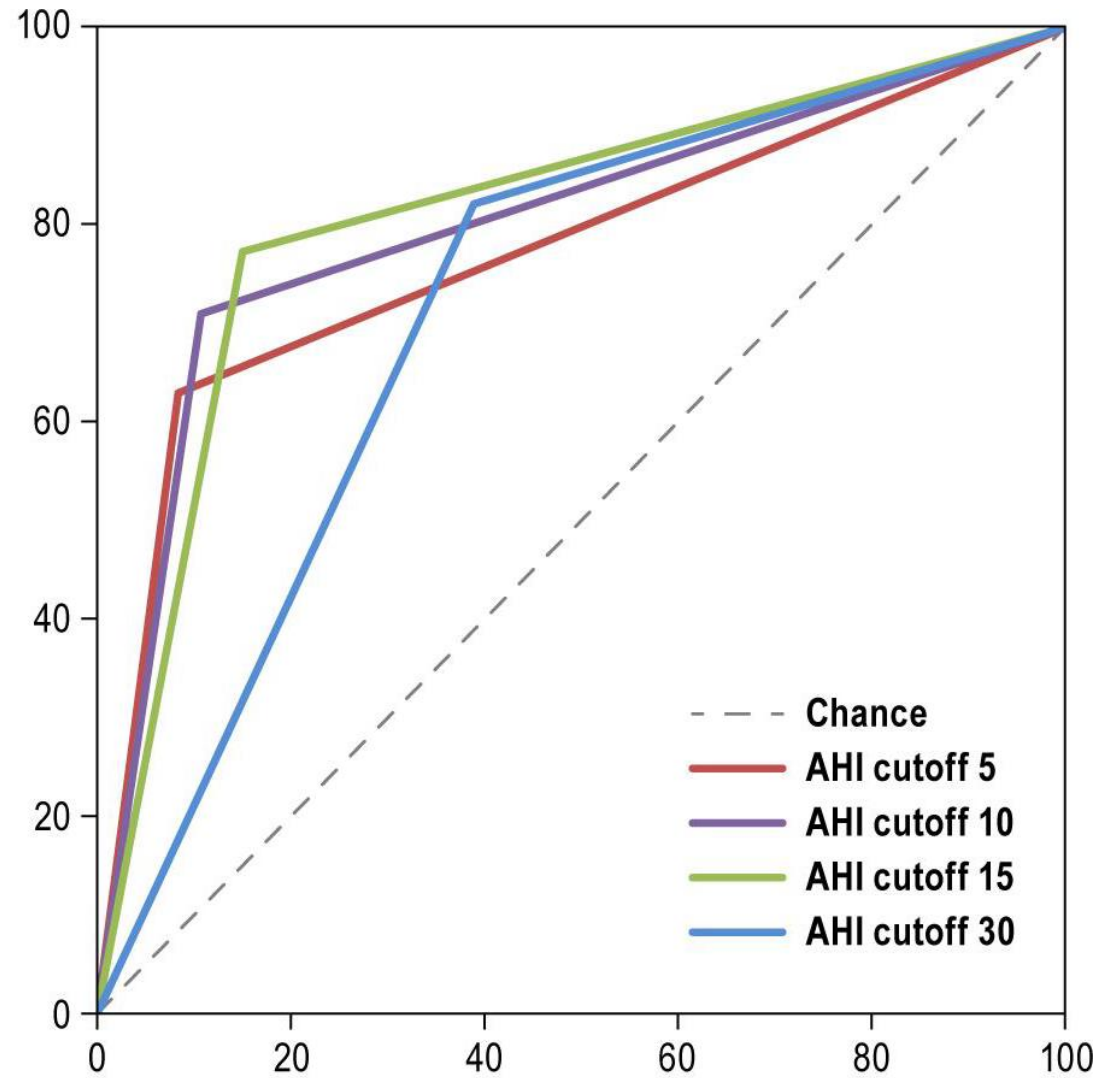
- Что если классификация всё же не вероятностная?



# ROC-AUC

- Что если классификация всё же не вероятностная?
- Строим кривую по **трем** точкам

# ROC-AUC



### 3. Пример



# Выбираем, что оптимизировать (на примере рекомендаций)

Что можем делать

- Прогнозировать, какие товары будут куплены
- Максимизировать прибыль
- Остается вопрос: какие прогнозы нужны и как их использовать, чтобы денег стало больше?

# Максимизация количества покупок

<b>Товар 1</b>	<b>Товар 2</b>	<b>Товар 3</b>	<b>Товар 4</b>
----------------	----------------	----------------	----------------

# Максимизация количества покупок

**Товар 1**

**Товар 2**

**Товар 3**

**Товар 4**

Вероятность

$p_1$

$p_2$

$p_3$

$p_4$

# Максимизация дохода

Товар 1	Товар 2	Товар 3	Товар 4
---------	---------	---------	---------

Вероятность	$p_1$	$p_2$	$p_3$	$p_4$
Цена	$c_1$	$c_2$	$c_3$	$c_4$

# Максимизация дохода



Puma  
Ветровка  
3 490 руб.



Crocs  
Сланцы  
1 990 руб.



Tony-p  
Слипоны  
~~1 999 руб.~~ 1 590 руб.



Champion  
Брюки спортивные  
~~3 599 руб.~~ 1 970 руб.

Вероятность	0.05	0.02	0.015	0.009
Цена	3490	1990	1590	1970



# Максимизация прибыли



Puma  
Ветровка  
3 490 руб.



Crocs  
Сланцы  
1 990 руб.



Tony-p  
Слипоны  
~~1 999 руб.~~ 1 590 руб.



Champion  
Брюки спортивные  
~~3 599 руб.~~ 1 970 руб.

Вероятность	0.05	0.02	0.015	0.009
Цена	3490	1990	1590	1970
Маржа	0.1	0.4	0.4	0.2

## Мини-задача

Как изменится построение модели, если нам нужно максимизировать количество просмотренных пользователем товаров?