

# Лекция 2

Описание задачи



он нас всех  
посчитал



## **«Нас всех посчитали»**

# **«Нас всех посчитали» или скоринговые системы**

Один из способов повышения эффективности взаимодействия банка с клиентами заключается в том, чтобы отправлять предложение о новой услуге не всем клиентам банка, а только некоторой части, выбираемой по принципу наибольшей склонности к отклику на данное предложение.

Конкурсное задание заключается в том, чтобы предложить алгоритм, который будет выдавать оценку склонности клиента к положительному отклику по его признаковому описанию. Эта оценка может (хотя и не обязана) интерпретироваться как вероятность положительного отклика. Предполагается, что, получив такие оценки для некоторого множества клиентов, банк обратится с предложением только к тем клиентам, у которых значение оценки выше некоторого порога.

- Исходная выборка («выборка А») содержит записи о 15 223 клиентов, классифицированных на два класса:
  - 1 — отклик был (1812 клиентов),
  - 0 — отклика не было (13411 клиентов).

**Credit** - обучающая выборка

- Ещё 14 910 записей отложены в качестве тестовых («выборка Б») – ответы по ним известны только банку.

**Credit\_new** - тестовая выборка

- Тестовые данные будут использоваться для определения победителя конкурса.

Записи (признаковые описания) клиентов состоят из 50 признаков, в состав которых входит, в частности, возраст, пол, социальный статус относительно работы, социальный статус относительно пенсии, количество детей, количество иждивенцев, образование, семейное положение, отрасль работы.



Качество работы алгоритма будет оцениваться, согласно сложившейся в банковском маркетинге практике, при помощи показателя **AUC** (area under curve).

$X$  – множество объектов

$Y$  – множество ответов

$Y'$  – множество ответов

$$X = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_k^{(1)} \\ x_1^{(2)} & \ddots & & \vdots \\ \vdots & & & \\ x_1^{(n)} & \dots & & x_k^{(n)} \end{pmatrix}$$

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$



$$Y' = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix},$$

**Найти:**

$a : X \rightarrow Y$  – решающую функцию (алгоритм), приближающую  $y$

**ПОЕХАЛИ!**