

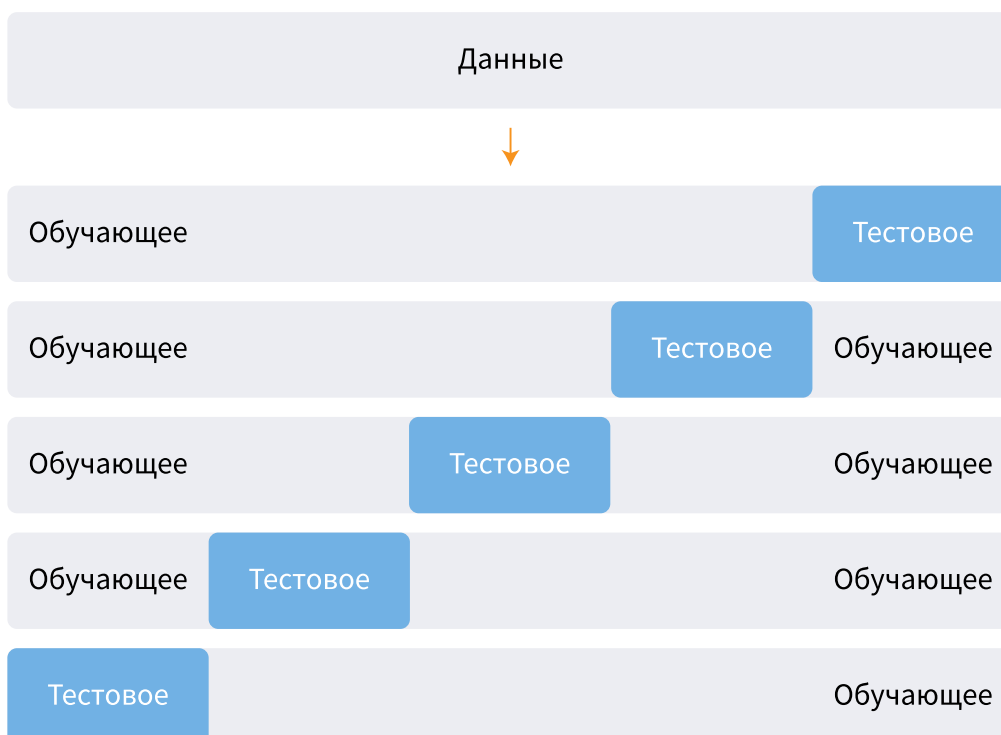
# Кросс-валидация (Cross-validation)

Синонимы: Перекрестная проверка, Rotation estimation

Loginom: [Валидация моделей](#)

Кросс-валидация — это метод оценки аналитической модели и ее поведения на независимых данных с наиболее равномерным использованием имеющихся данных.

В основе метода лежит разделение исходного множества данных на  $k$  примерно равных блоков, например  $k = 5$ . Затем на  $k - 1$ , т.е. на 4-х блоках, производится обучение модели, а 5-й блок используется для тестирования. Процедура повторяется  $k$  раз, при этом на каждом проходе для проверки выбирается новый блок, а обучение производится на оставшихся.



Перекрестная проверка имеет важные преимущества перед применением одного множества для обучения и одного для тестирования модели: если при каждом проходе оценить выходную ошибку модели и усреднить ее по всем проходам, то полученная ее оценка будет более достоверной.

На практике чаще всего выбирается  $k = 10$  (10-ти проходная перекрестная проверка), когда модель обучается на 9/10 данных и тестируется на 1/10. Исследования показали, что в этом случае получается наиболее достоверная оценка выходной ошибки модели.