

Во-первых, для чего нужно делить данные на тестовую и обучающую выборку?

Существуют такие проблемы, как переобучение (model-overfitting) и недообучение (model-underfitting)

## Проблема переобучения

С точки зрения данных проблему «заучивания билетов» еще называют переобучением модели (model overfitting). Когда мы учим и проверяем модель на одних и тех же данных, то по сути просим модель запомнить каждую единицу наших данных.

Также существует проблема недообучения (underfitting).

Если бы мы строили не линейную модель, а, например, полиномиальную (она выражена кривой), то нормально обученная модель отличалась бы от недообученной и переобученной следующим образом.

