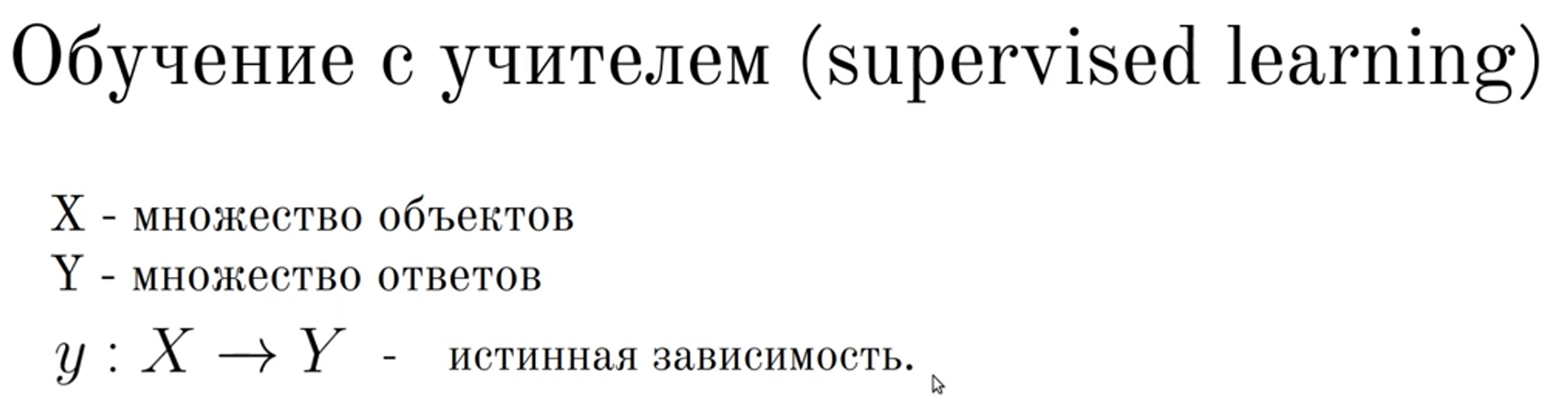
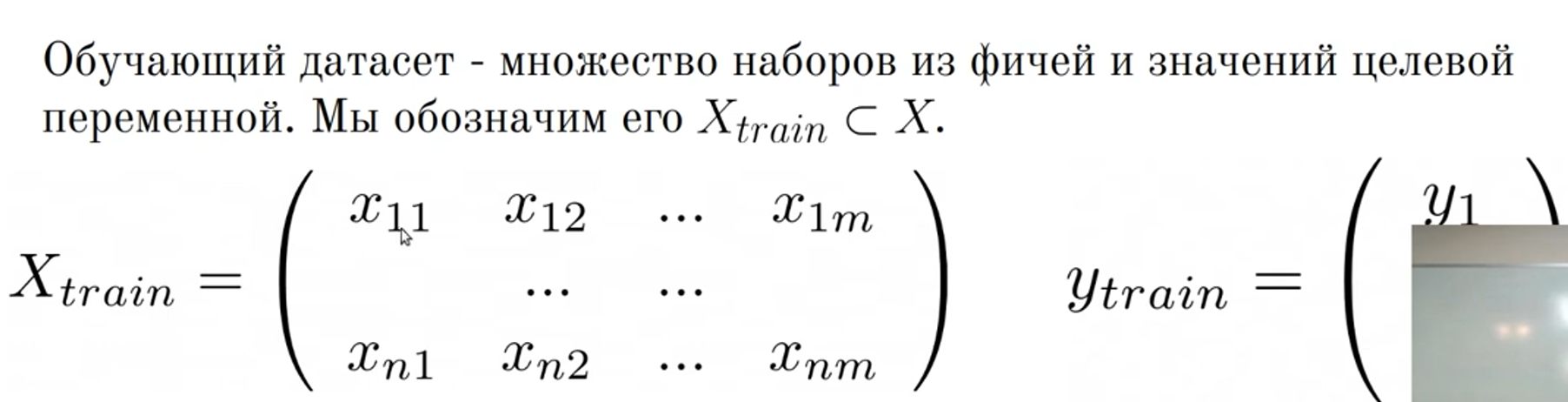
Введение в машинное обучение.



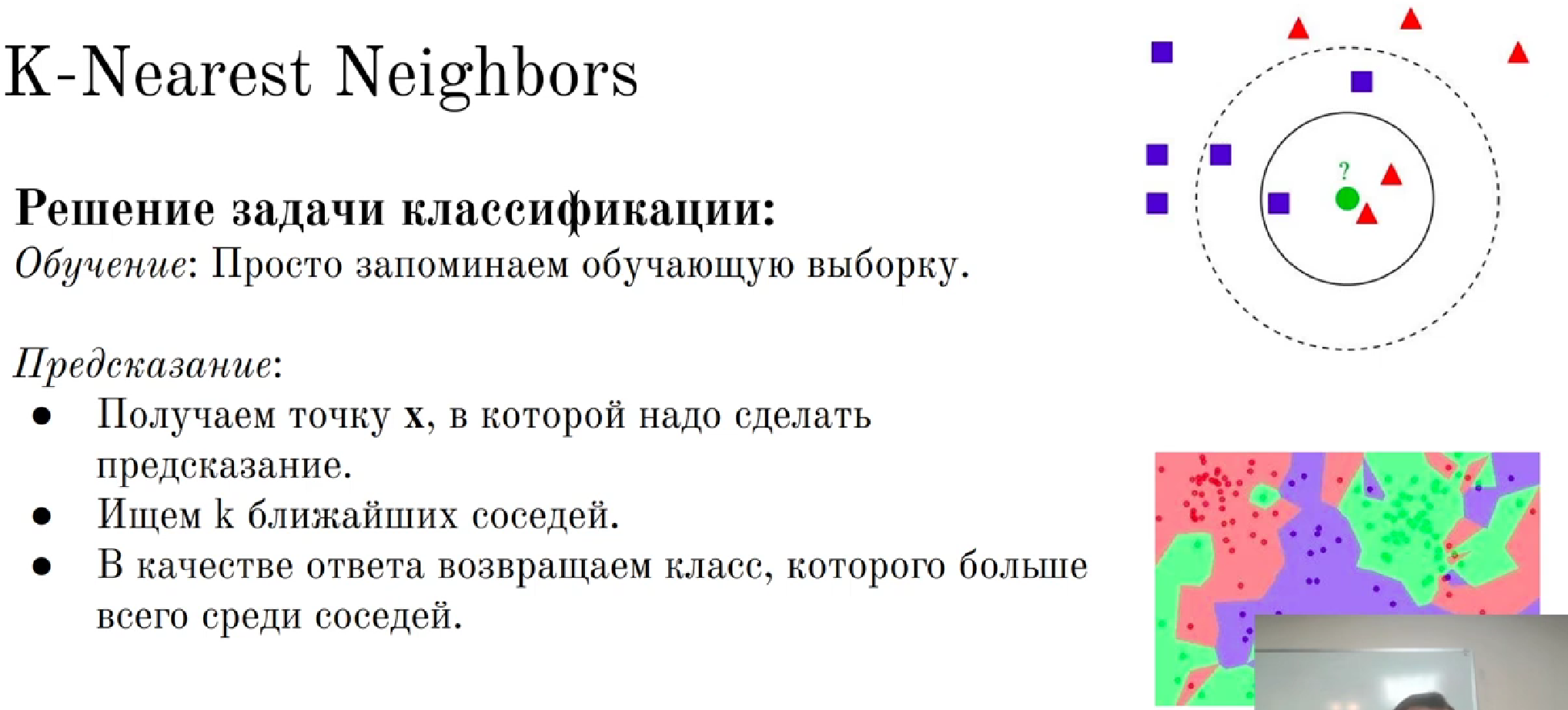
Функция y, которая всем объектам из Х сопоставляет объекты из множества Y. Такую функцию называют истинная зависимость, функция закона природы, функция, которую мы хотим восстановить, target, целевая переменная. Мы обучаем нейронную сеть и таким образом создаем эту самую функцию, которую потом сможем использовать в реальной жизни, на реальных объектах.

Обучение с учителем значит, что есть набор объектов и набор ответов к этим объектам. Путем анализа двух этих наборов мы можем восстановить эту самую функцию, т.е. обучить нейросеть.



Xtrain – матрица, состоящая из объектов, свойства которого были заданы в вектор. Каждая строчка – свойства для отдельного объекта. Компоненты вектора в датасете Xtrain у каждого вектора называются признаками или фичами.

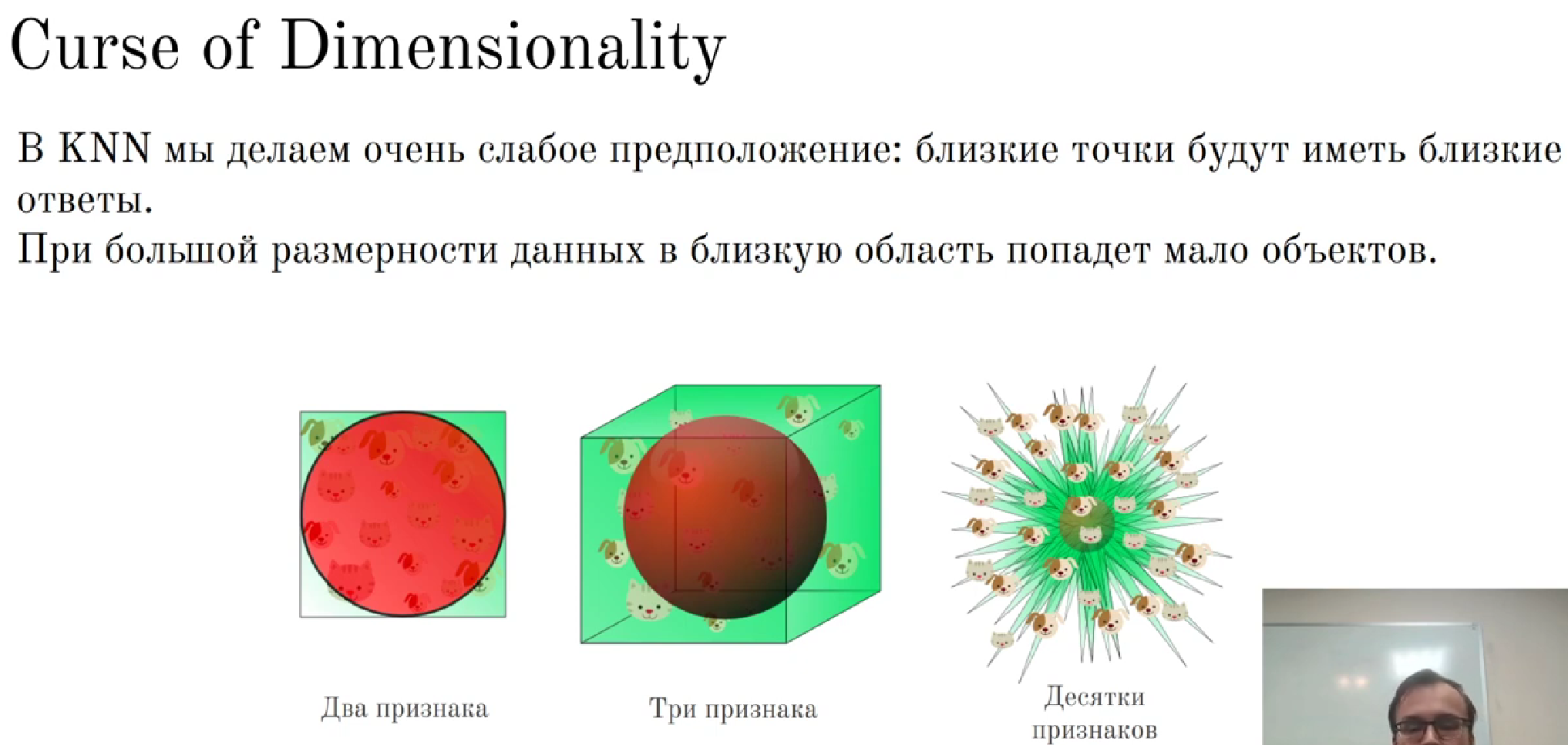
Xtrain – такую матрицу называют обучающей выборкой или просто выборкой.



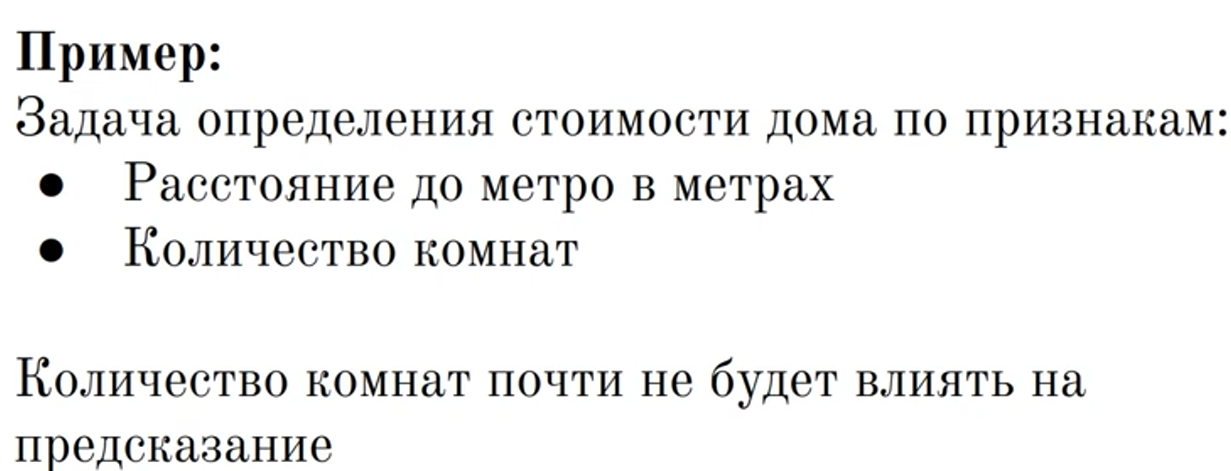
Мы запоминаем всю обучающую выборку и строим точки, которые соответствуют концам векторов из этой выборки, мы знаем цвет каждой точки. Предсказание происходит как написано на картинке.

k – это гиперпараметр, он так называется потому что не вычисляется во время обучение нейросети, мы сами вручную должны задать его. Насколько хорошо будет работать нейросеть фактически зависит от него.

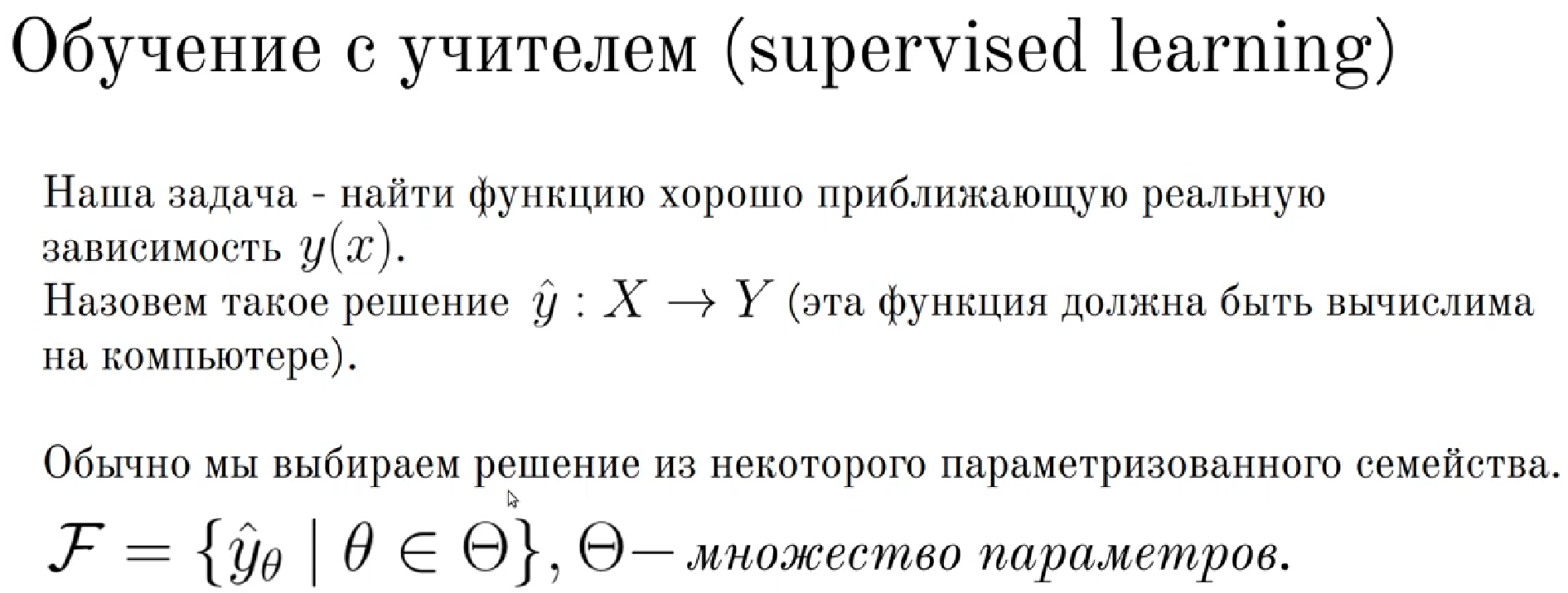
Эта модель работает на основе предположения, что если у двух объектов признаки более-менее близки, то и их ответы тоже должны быть схожи. Это предположение нарушается, когда наши объекты имеют много признаков, в таком случае близкие точки могут и не иметь близкие ответы изза большой многомерности пространства, в какой-то близкой окрестности может и вовсе не быть никаких точек. Такое явление называется проклятием размерности.



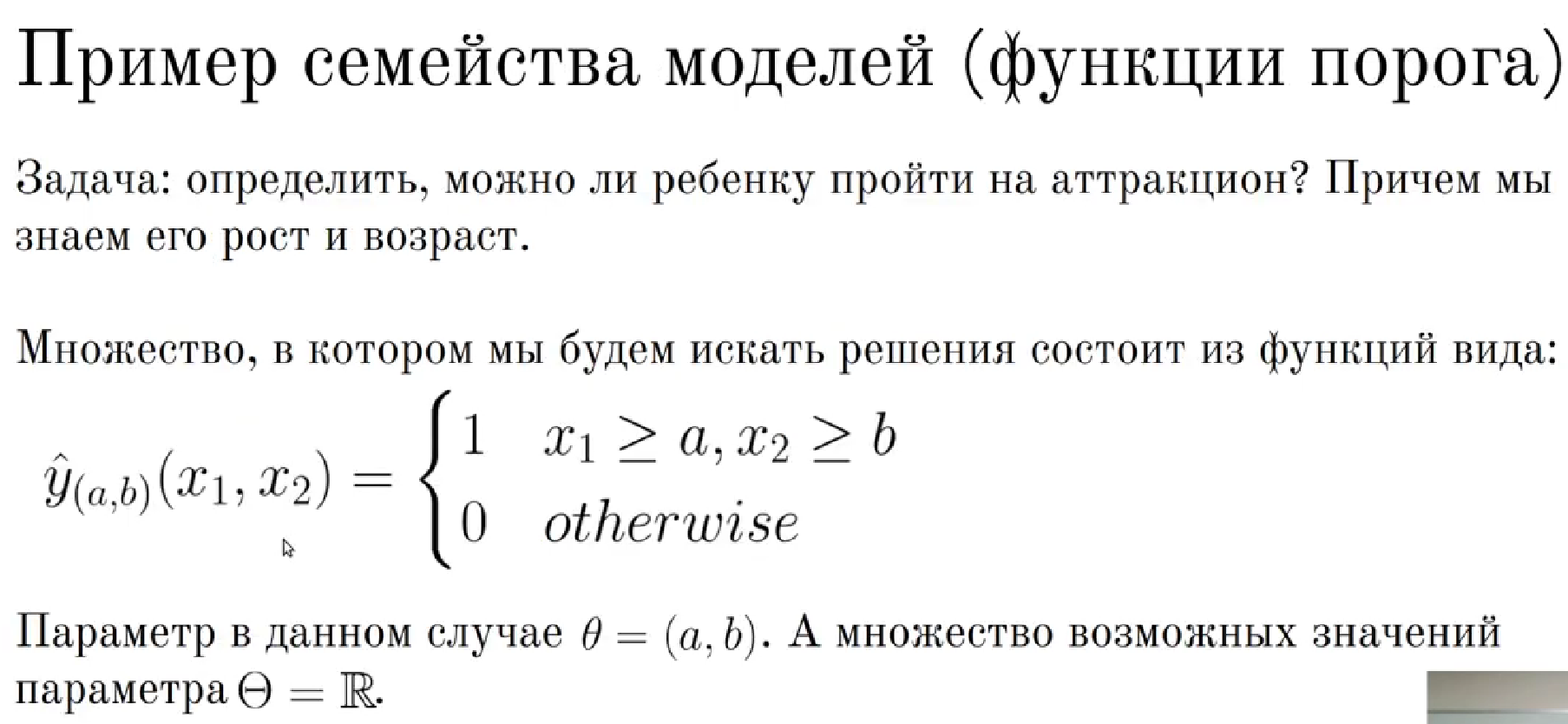
Также серьезной проблемой этой модели является масштаб признаков.



Расстояние до метро измеряется в сотнях-тысячах метрах, а количество комнат максимум в десятках штук. Но и тот и другой признак влияет на положение точки в пространстве. И получается, что количество комнат в квартире практически не будет влиять на стоимость квартиры, но расстояние до метро будет играть ключевую роль.

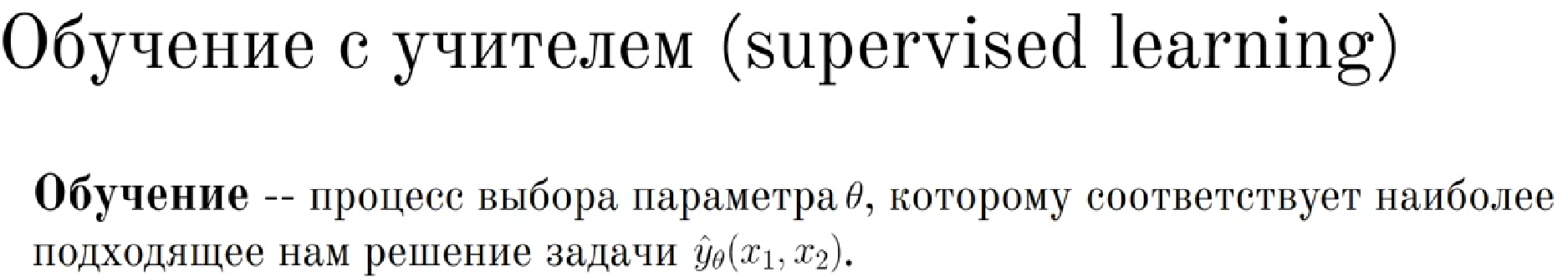


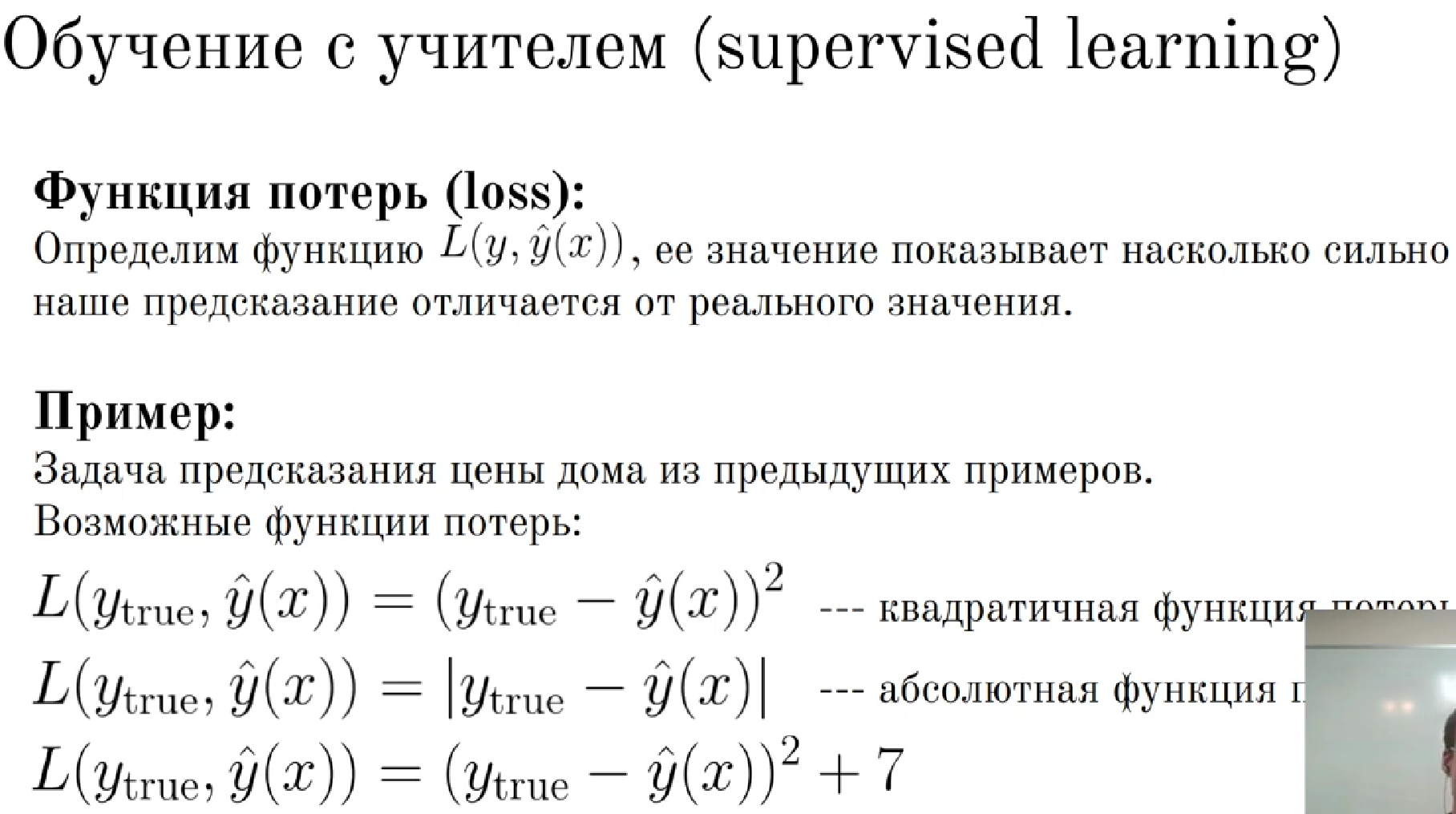
F – это параметризованное семейство. Т.е. такое множество функций, в котором каждой функции соответствует какой-то параметр. В зависимости от выбранного нами параметра, из этого параметризованного семейства будет выбрана соответствующая функция.

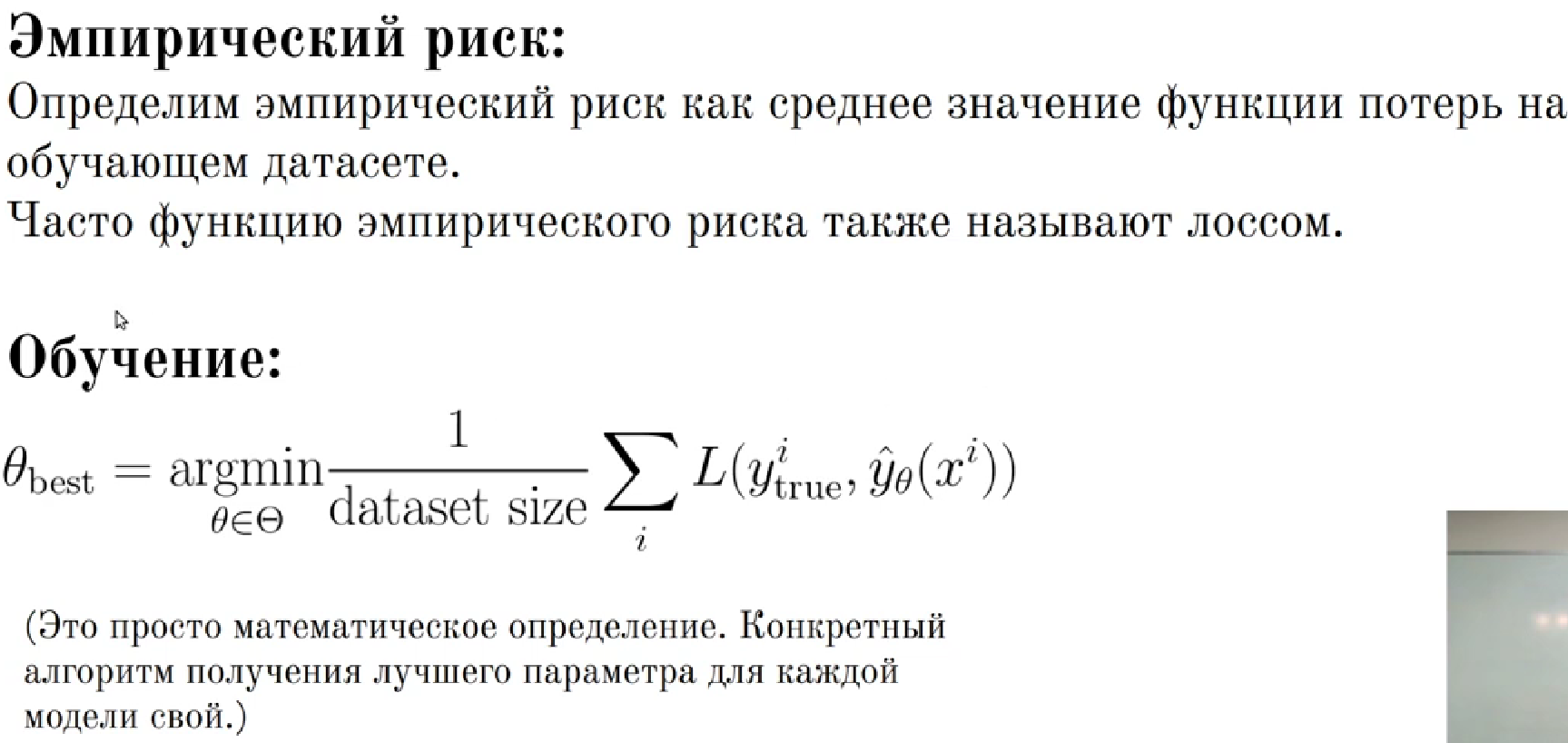


Множества возможных значений параметра равно R\*\*2, где R – количество признаков.

Таким образом, у нас есть некая функция порога, внутри которой есть параметры. В ходе обучения модели эти параметры изменяются так, чтобы функция выдавала результат, наиболее близкий к правильному. Таким образом обучение модели – это подбор функции, ну или точнее – подбор параметров для нее.

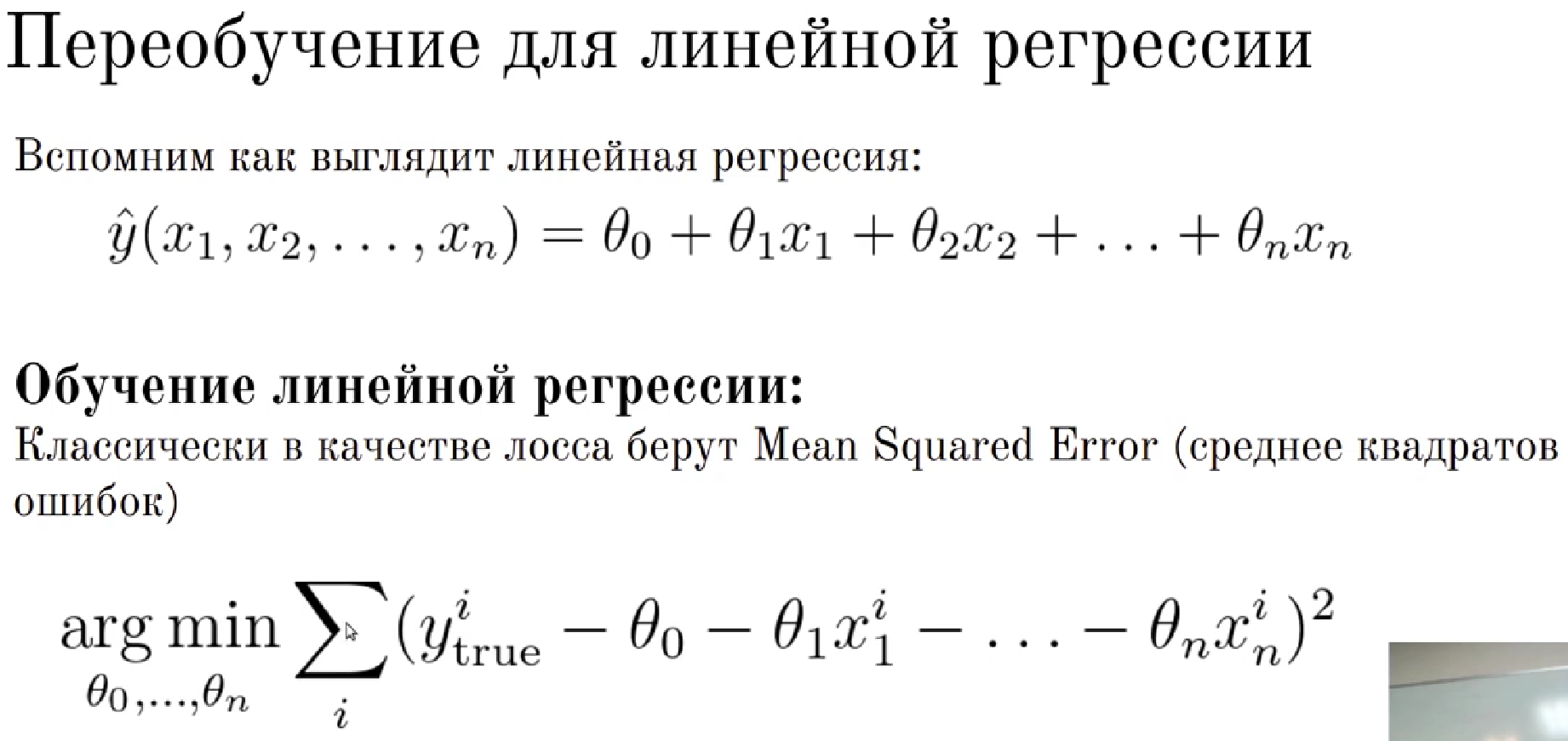


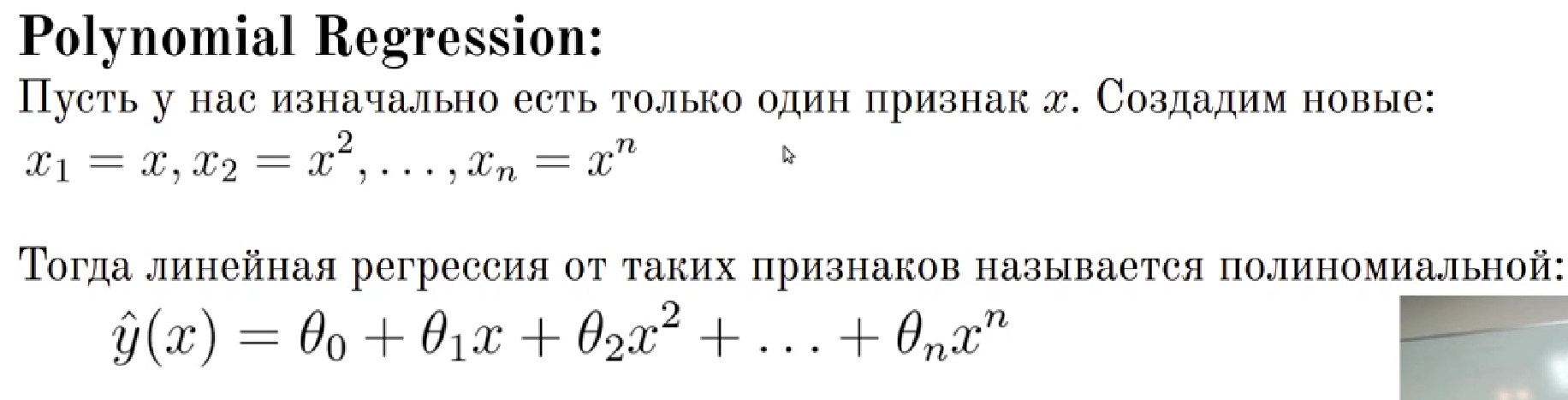


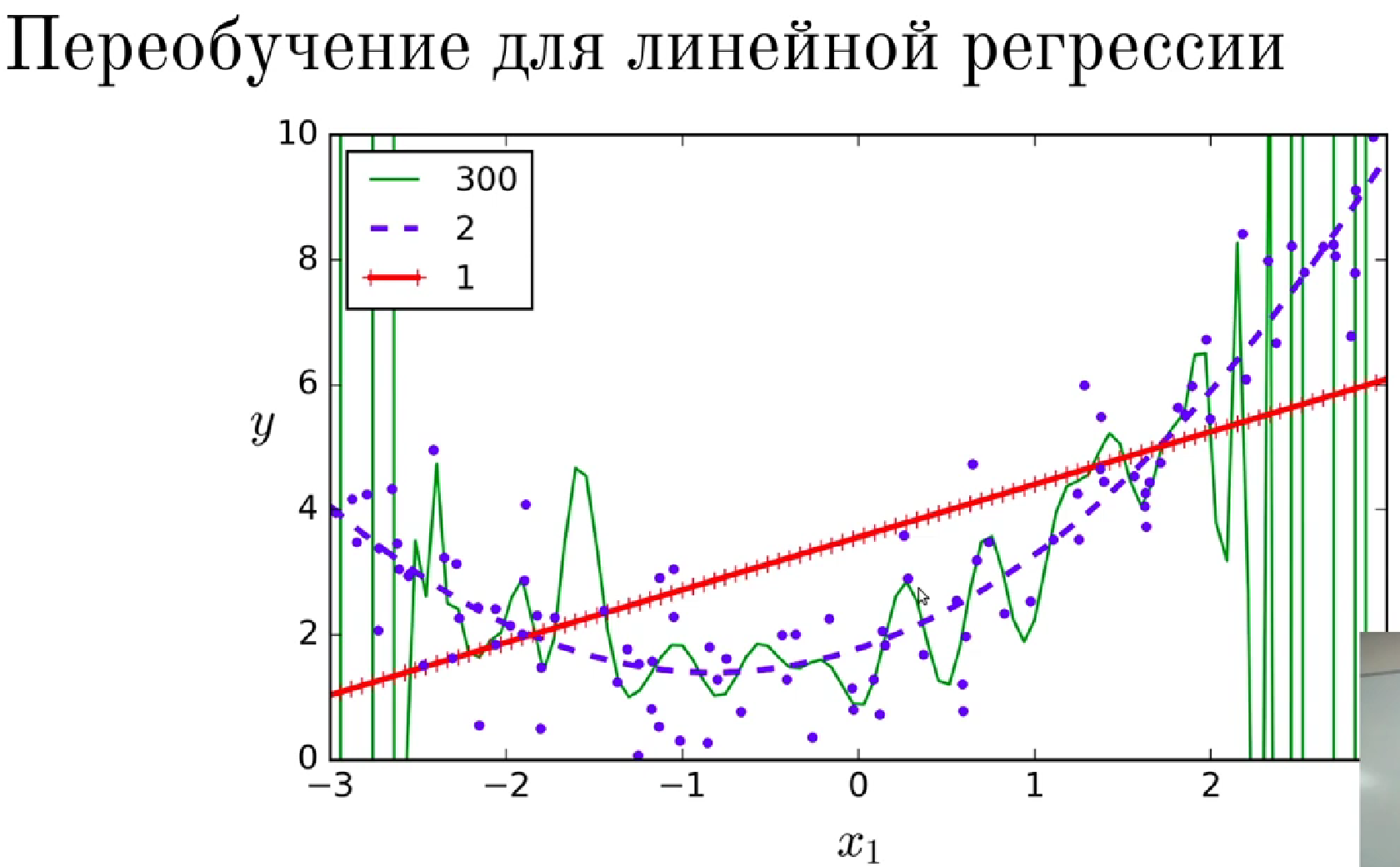


Функция argmin значит, что есть какая то функция (которая пишется справа от argmin), которая зависит от какого то параметра, которое пишут под argmin. И argmin возвращает этот параметр, который соответствует наименьшему значению функции.

Таким образом, функция, показанная на рисунке, нужна для минимализации эмпирического риска.







Если модель нашла какие то зависимости в обучающей выборке, которых на самом деле нет и в реальности начала выдавать плохие результаты, то это значит, что мы это модель переобучили.

Чтобы осознавать, что модель не переобучена, нужно следовать правилу создания моделей, вот эти этапы:

