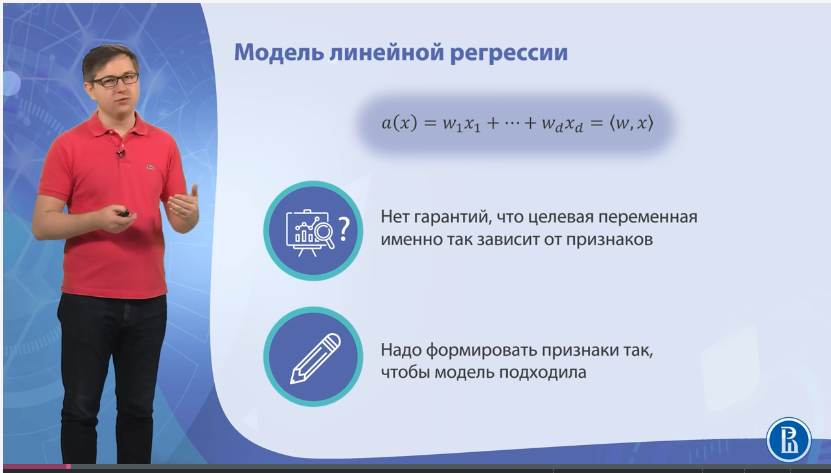
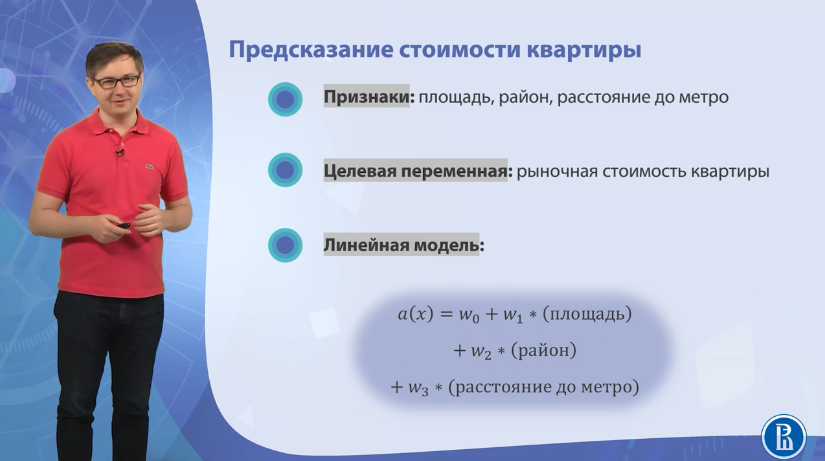
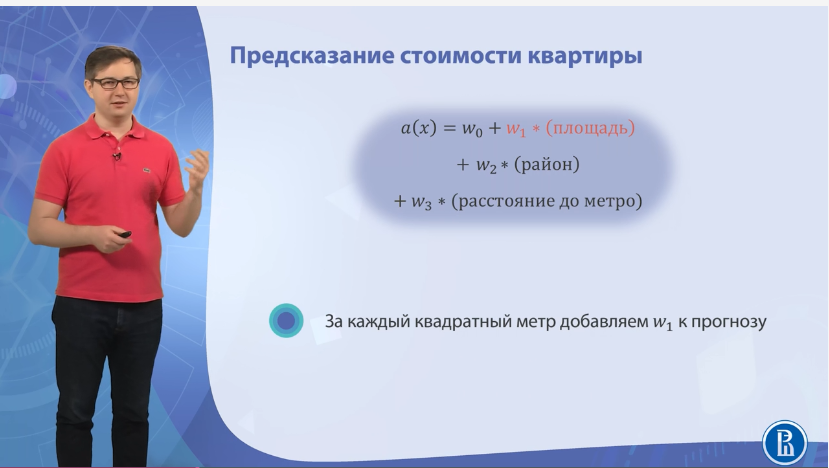


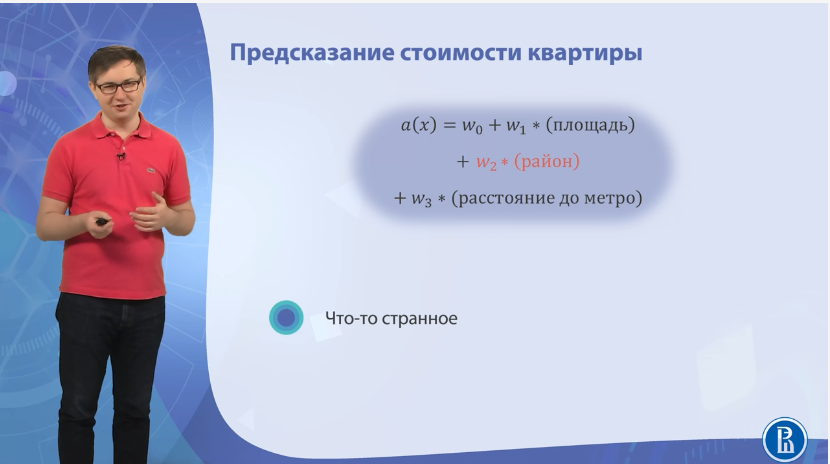
Модель - Сумма признаков с линейными коэффициентами

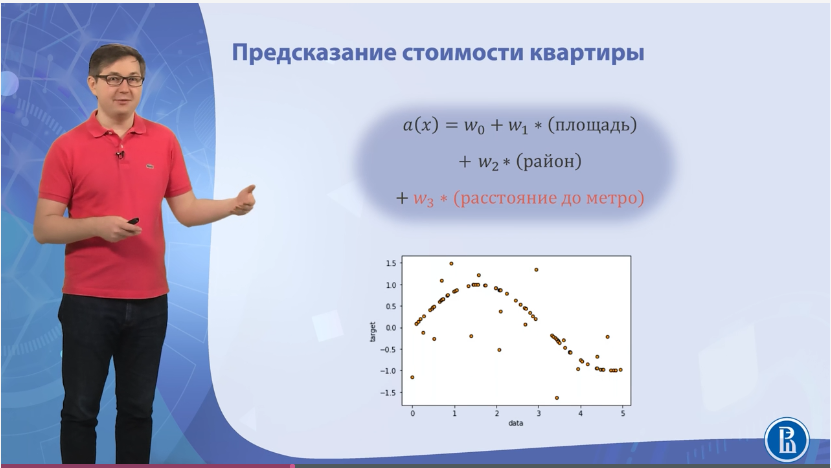




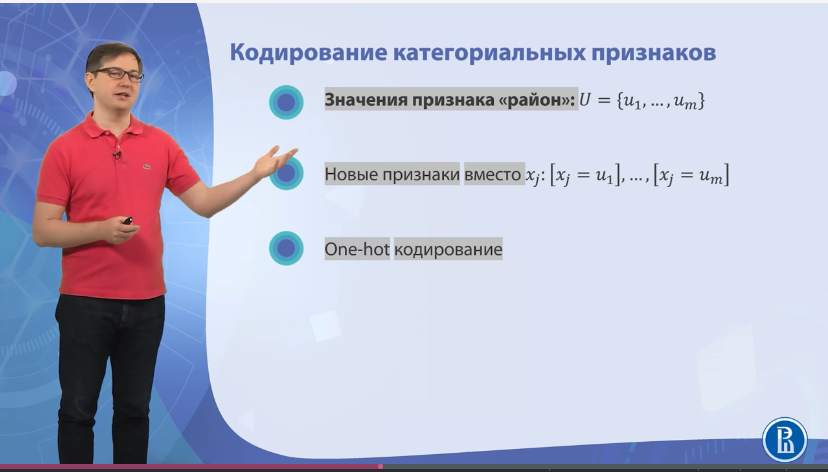
Рекомендательная система – клиент вводит данные по своей квартире, и система выводит рекомендацию по рыночной стоимости







Может быть нелинейная зависимость стоимости от расстояния до метро



2 признак – район – категориальный признак

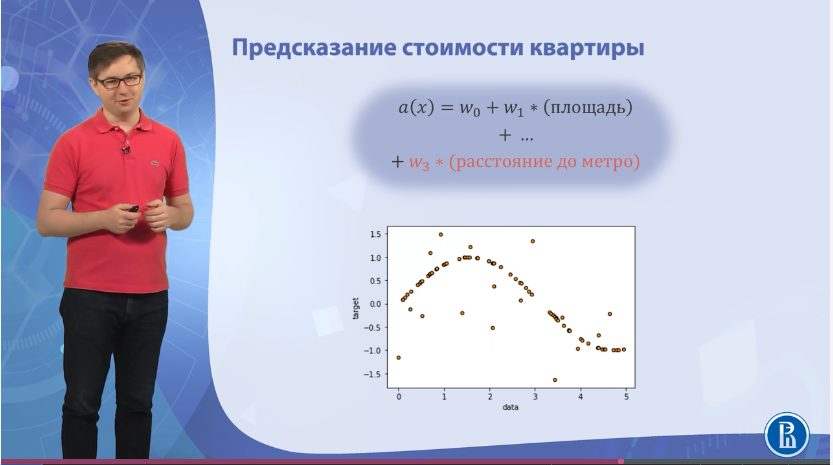
Есть способы кодирования категориального признака, самый популярный – бинарное кодирование или One-hot. Вместо одного признака заводим множество бинарных признаков, т.е равен ли признак определенному значению. Т.е. выделяем m бит на кодирование.

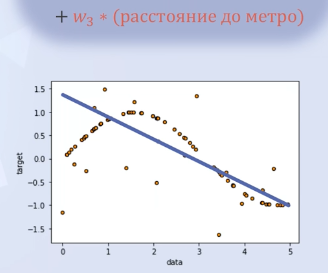


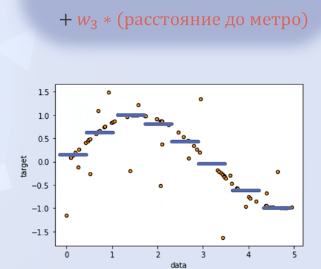
Это пример



Получится, что w2 – добавка ко всем квартирам, находящимся в ЦАО, w3 – добавка ко всем квартирам, находящимся в ЮАО, и т.д.



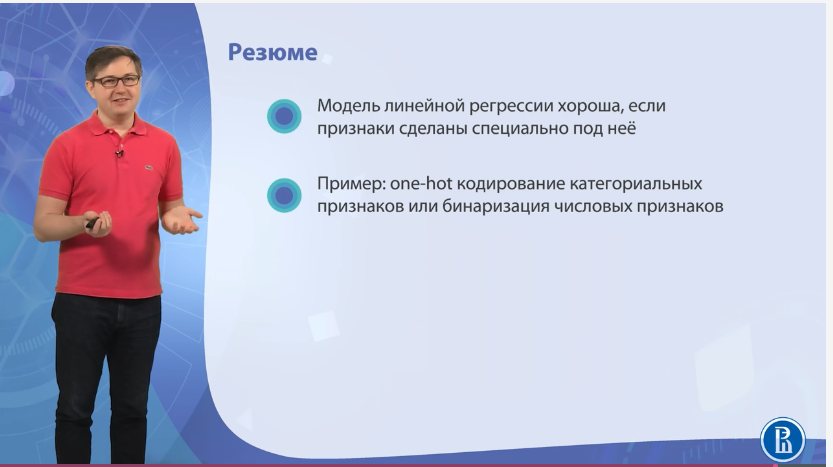
 - плохой прогноз

 - желаемый прогноз

Надо также заменить один признак на множество других



Бьем весь интервал значений по X на минимально допустимые отрезки и объявим индикаторы попадания в каждый отрезок новыми признаками. Каждый признак умножим на свой коэффициент



**Задания на понимание**

 Добавить страницу в мои закладки

Давайте проверим, насколько хорошо вы поняли материал прослушанной лекции.

1 из 1 балла (не оценивается)

Рассмотрим признак “Образовательная программа” при анализе данных по студентам университета. Этот признак может принимать три значения: “Экономика”, “Математика”, “Философия”. Воспользуемся one-hot кодированием и заменим этот признак на три бинарных, которые будут соответствовать категориям в том порядке, в котором они перечислены выше. Как будет закодирован признак со значением “Философия”?

(0, 0, 1) верно

(0, 1, 0)

(1, 0, 0)