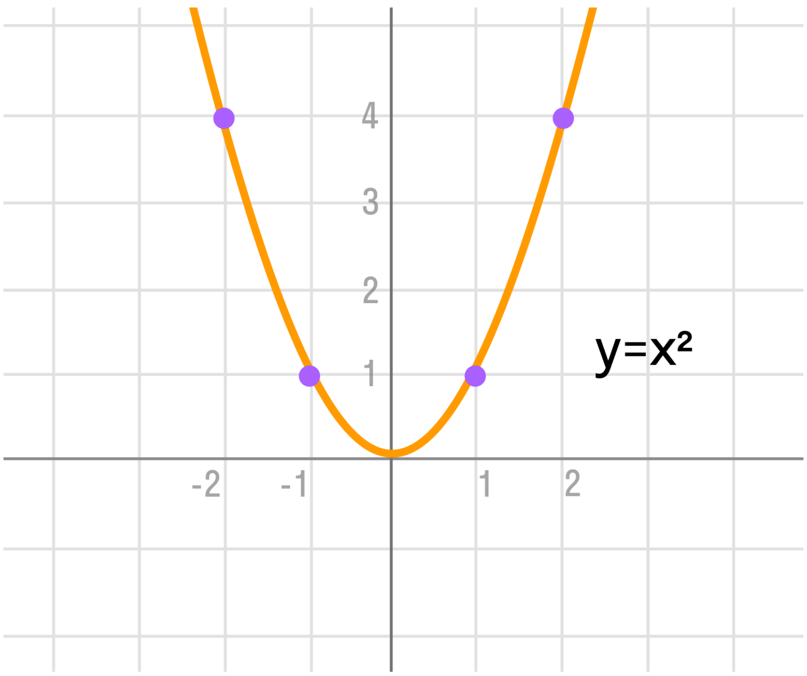
2. Оценка качества работы моделей

telegram: @ashagvaleev

Ликбез №1: Понятие функции функциональной зависимости

Функция — это некоторое правило, по которому одно множество связано с другим.

Пример функциональной зависимости, где множество X отображается во множестве Y является формула $y = x^2$



В МО мы также сталкиваемся с необходимостью сопоставления множества одних объектов с другим через определенные функциональные зависимости.

Композиция функций — когда аргументом одной функции является другая функция.

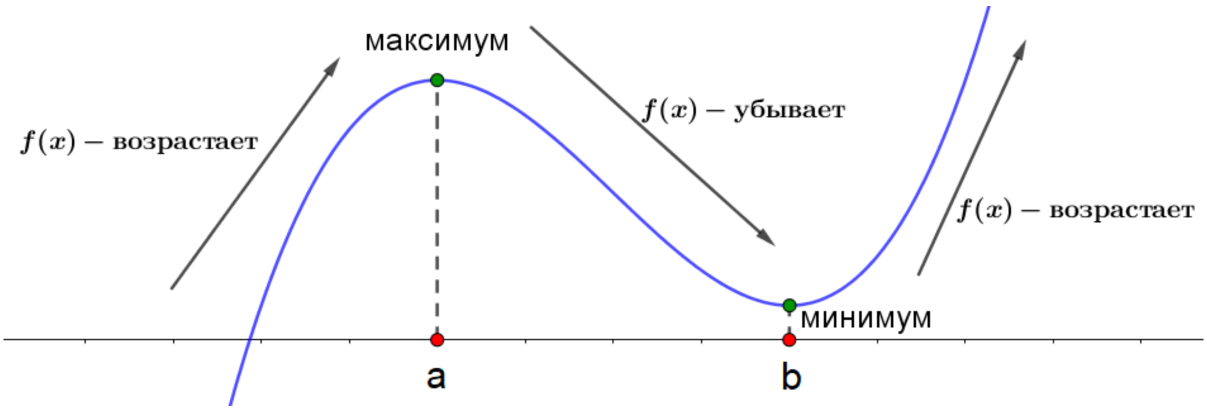
Например, в качестве аргумента (x) берем функцию синуса $f(x) = sin(x)$ и возводим её в квадрат:

$$z(x) = g(f(x)) = f(x)^2 = (sin(x))^2$$

Чаще всего в МО мы имеем дело с композициями функций.

Минимумы и максимумы функции.

Минимум (максимум) функции — точка (аргумент), в котором значение функции минимально (максимально) в ближайшей окрестности. Локальных минимумов (максимумов) может быть несколько, глобальный только один.



Функция потерь (Loss function)

Допустим, у нас есть некая модель $a(x)$, определяющая зависимость для пар объект-таргет из выборки

$$X = (x_i, y_i)$$

Чтобы понять, хороша ли данная модель, оценим ошибку на одном из объектов. Для этого нам понадобится функция потерь.

Функция потерь — это ошибка алгоритма на выбранном объекте. Показывает, насколько далек результат, полученный выбранным алгоритмом, от ожидаемого результата.

Одна из самых популярных функций потерь — квадратичное отклонение:

$$L(a(x_i), y_i) = (a(x_i) - y_i)^2$$

В качестве примера допустим, что нам нужно предсказать ответ - цену акции через год по признакам (конкуренты, капитал, текущая цена). Предположим, что год прошел и нам уже известны реальные значения таргета. У нас есть два объекта:

$$x_1 = (21, 165, 58),\ y_i = 45$$
$$x_2 = (45, 189, 101),\ y_i = 36$$

Также мы имеем оценки, полученные моделью:

$$a(x_1) = 46$$
$$a(x_2) = 33$$

Тогда мы можем посчитать потери на каждом объекте:

$$L(x_1, y_1) = (46 - 45)^2 = 1$$
$$L(x_2, y_2) = (33 - 36)^2 = 9$$

Функционал качества и метрики

Если усреднить функцию потерь по всем объектам, то получим среднюю потерю работы нашей модели $a(x)$ на выборке X из m объектов:

$$Q(a(x), X) = \frac{1}{N} \sum_{i=1}^{m} L(a(x_i), y_i)$$

Такая функция позволяет оценить среднюю ошибку модели по всей выборке, а не на одном конкретном объекте и будет в данном случае функционалом качества нашей модели.

Метрика — критерий, по которому мы окончательно замеряем качество модели. В некоторых случаях может совпадать с функционалом качества. **Метрика** - показатель, более приближенный к бизнес-задачам.

В дальнейшем мы будем говорить о метриках, когда будем разбирать конкретные методы машинного обучения.

Посчитаем функционал качества для предыдущего примера с предсказанием стоимости акций, в котором:

$$y_1 = 45, nbsp; nbsp; a(x_1) = 46$$
$$y_2 = 36, nbsp; nbsp; a(x_2) = 33$$

$$L(x_1, y_1) = (46 - 45)^2 = 1$$
$$L(x_2, y_2) = (33 - 36)^2 = 9$$

Подставив значения в формулу, получаем функционал качества:

$$Q = \frac{1}{2}(9 + 1) = 5$$

Полученное значение называют среднеквадратичным отклонением. В разных задачах разные метрики, важно найти адекватную метрику для выбранной модели.

Главные метрики в задачах регрессии

Средняя квадратическая ошибка (*MSE - Mean Squared Error*) - самый простой и распространенный показатель для оценки регрессии, определяется уравнением:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(a(x_i) - y_i)^2$$

где y_i фактический результат, а $a(x_i)$ прогноз модели. Сильно штрафует за сильные отклонения.

Средняя абсолютная ошибка(*MAE - Mean Absolute Error*) - ошибка рассчитывается как среднее абсолютных разностей между целевыми значениями и прогнозами. Основным преимуществом MAE является то, что она является более устойчивой к выбросам, так как модули так не увеличивают отклонения, как возведение в степень. Одинаково штрафует как за сильные, так и за слабые отклонения.

$$MSE = \frac{1}{N}\sum_{n=1}^{N}|a(x_i) - y_i|$$

Полезная [статья](#) о применении данных метрика в задачах регрессии.