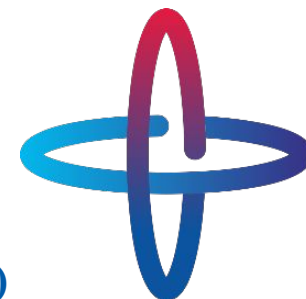


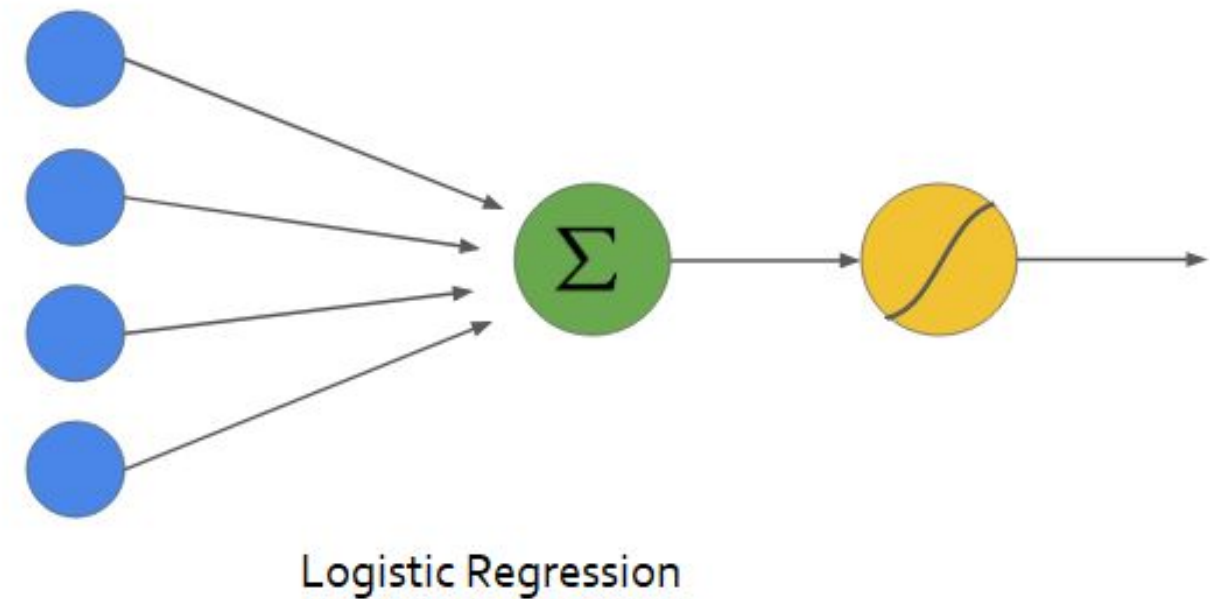
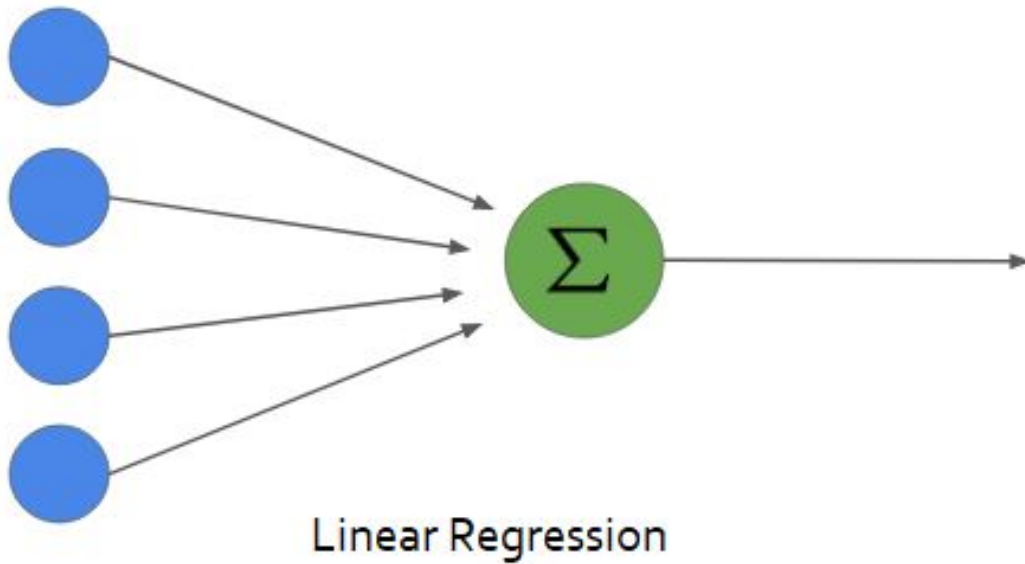
Системы искусственного интеллекта

Логистическая регрессия

Анна Герасимец (1842)
Яков Вишневский (1842)



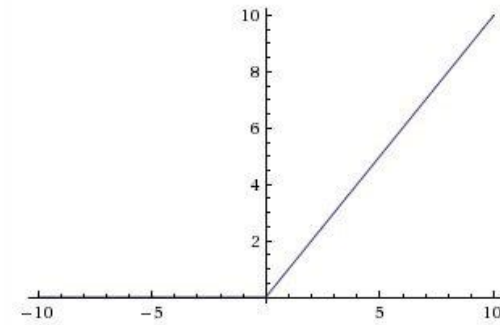
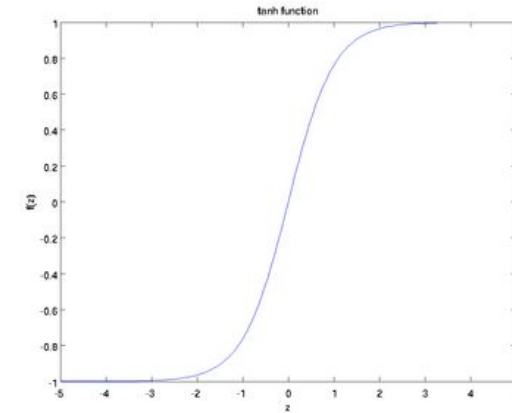
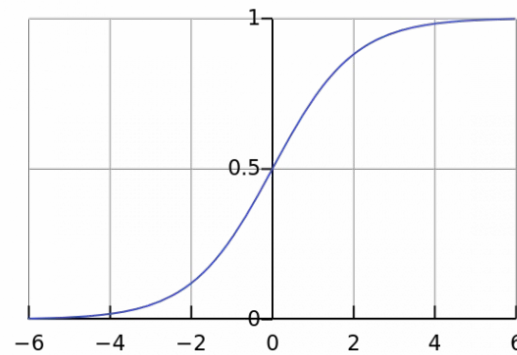
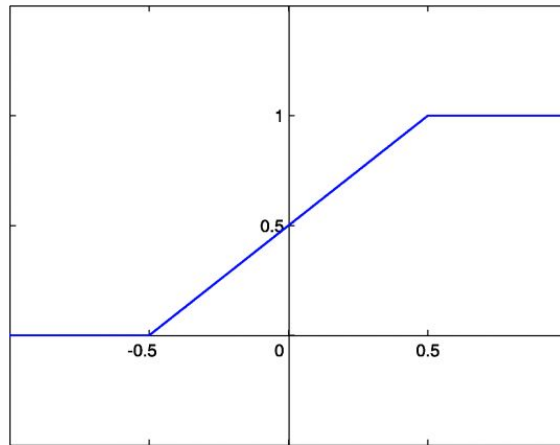
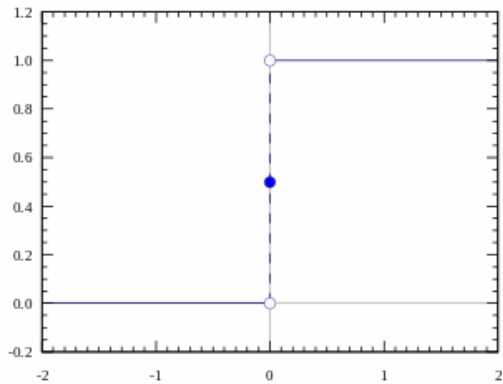
ГУАП



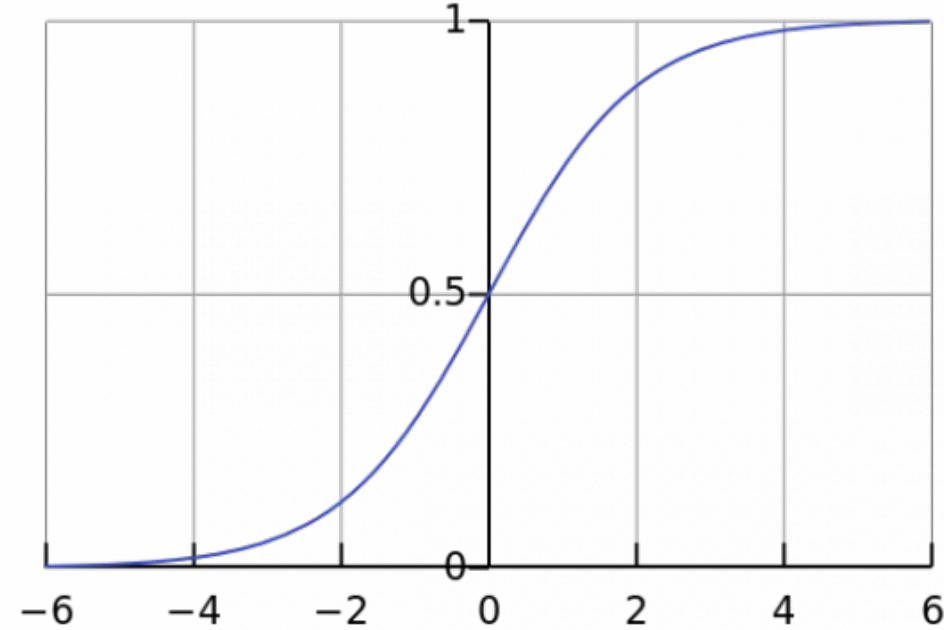
Логистическая регрессия вычисляет вероятность того, что данное исходное значение принадлежит к определенному классу. Она используется для задач классификации.

Функция активации определяет выходное значение нейрона в зависимости от результата взвешенной суммы входов и порогового значения.

Что делает **искусственный нейрон**? Он считает взвешенную сумму на своих входах, добавляет смещение (bias) и решает, следует это значение исключать или использовать дальше



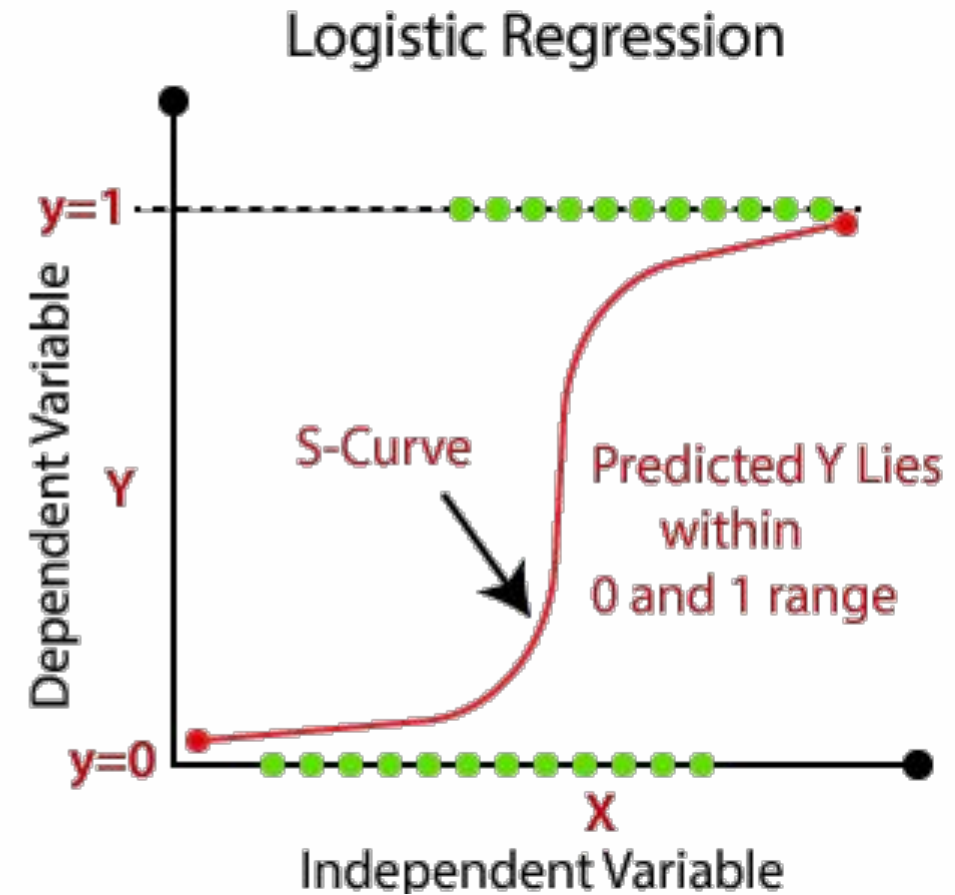
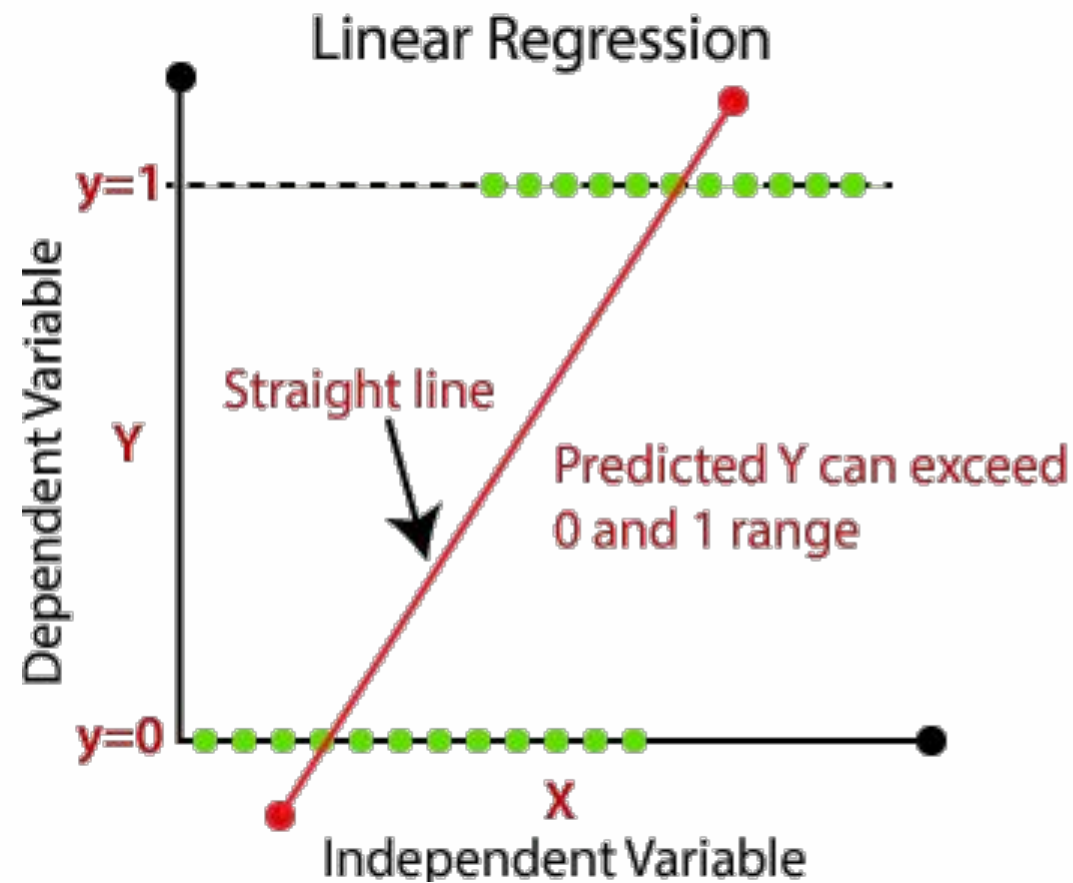
$$A = \frac{1}{1+e^{-x}}$$

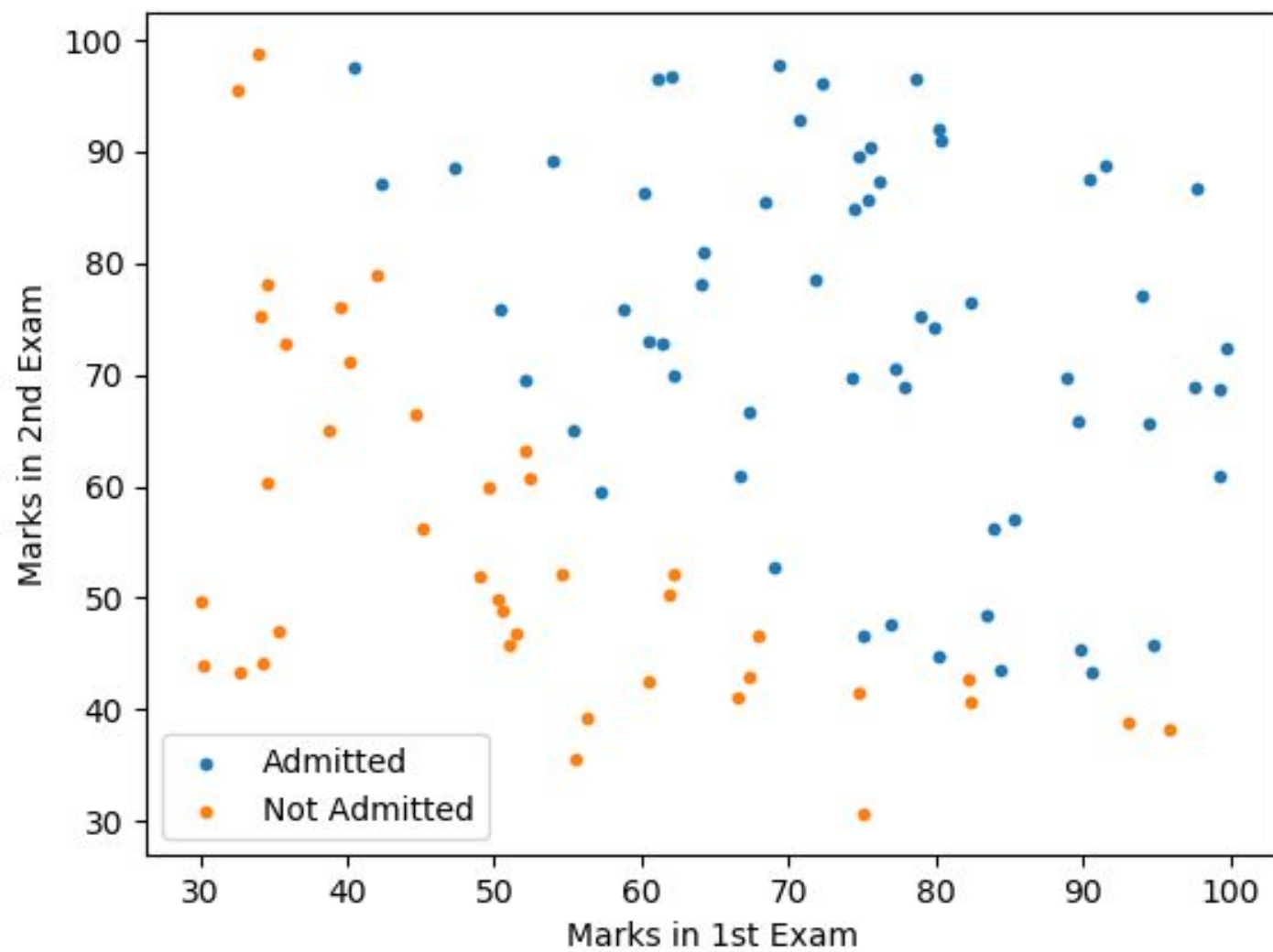


Сигмоида подходит для задач классификации:

- стремиться привести значения к одной из сторон кривой
- имеет фиксированный диапазон значений функции — $[0,1]$

Такое поведение позволяет находить четкие границы при предсказании и не приводит к ошибкам в случае больших значений активации.





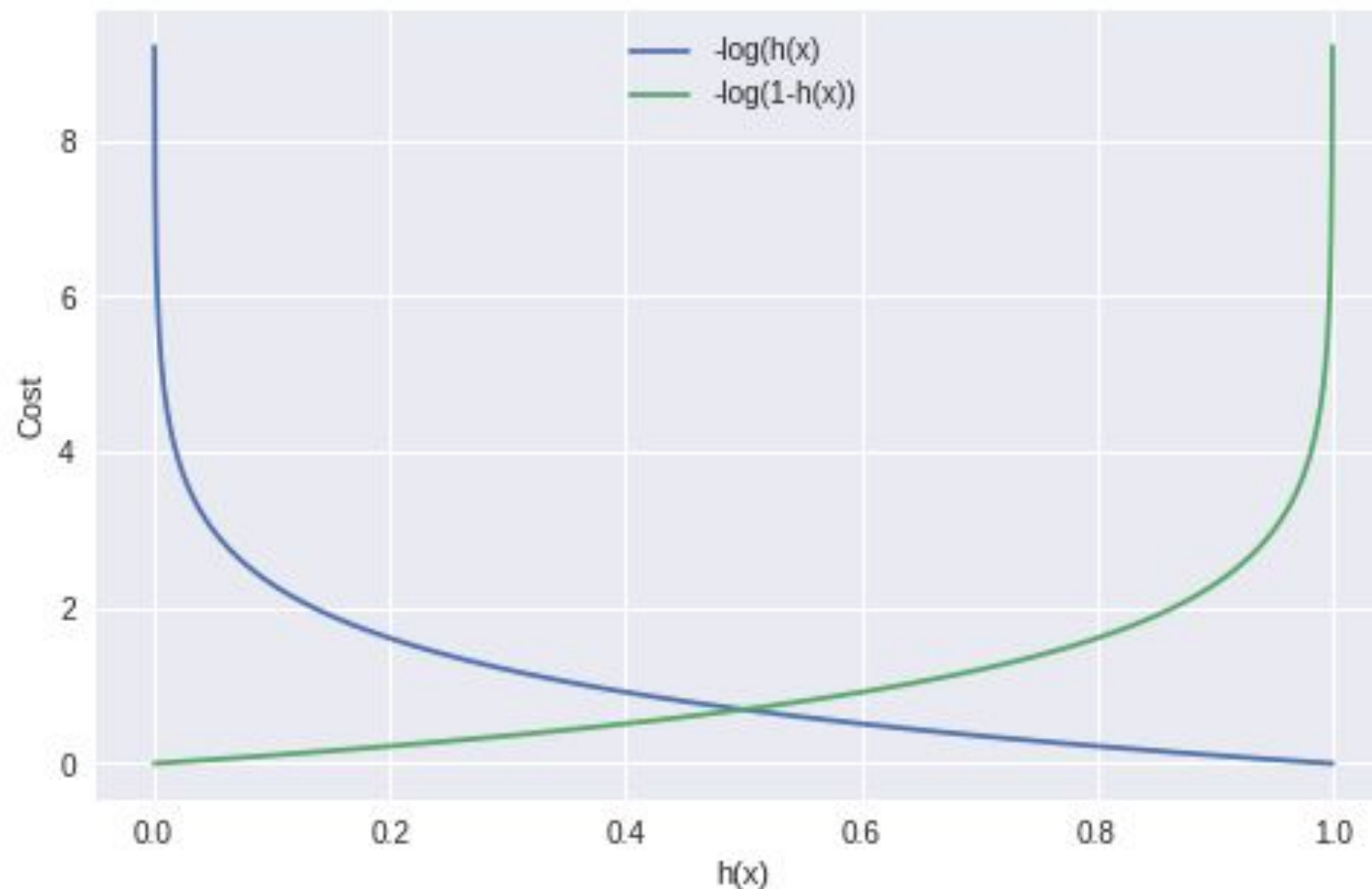
$$h(x) = \theta^T x$$

$$h(x) = \sigma(\theta^T x)$$

$$\sigma(t) = \frac{1}{1 + e^{-t}}$$

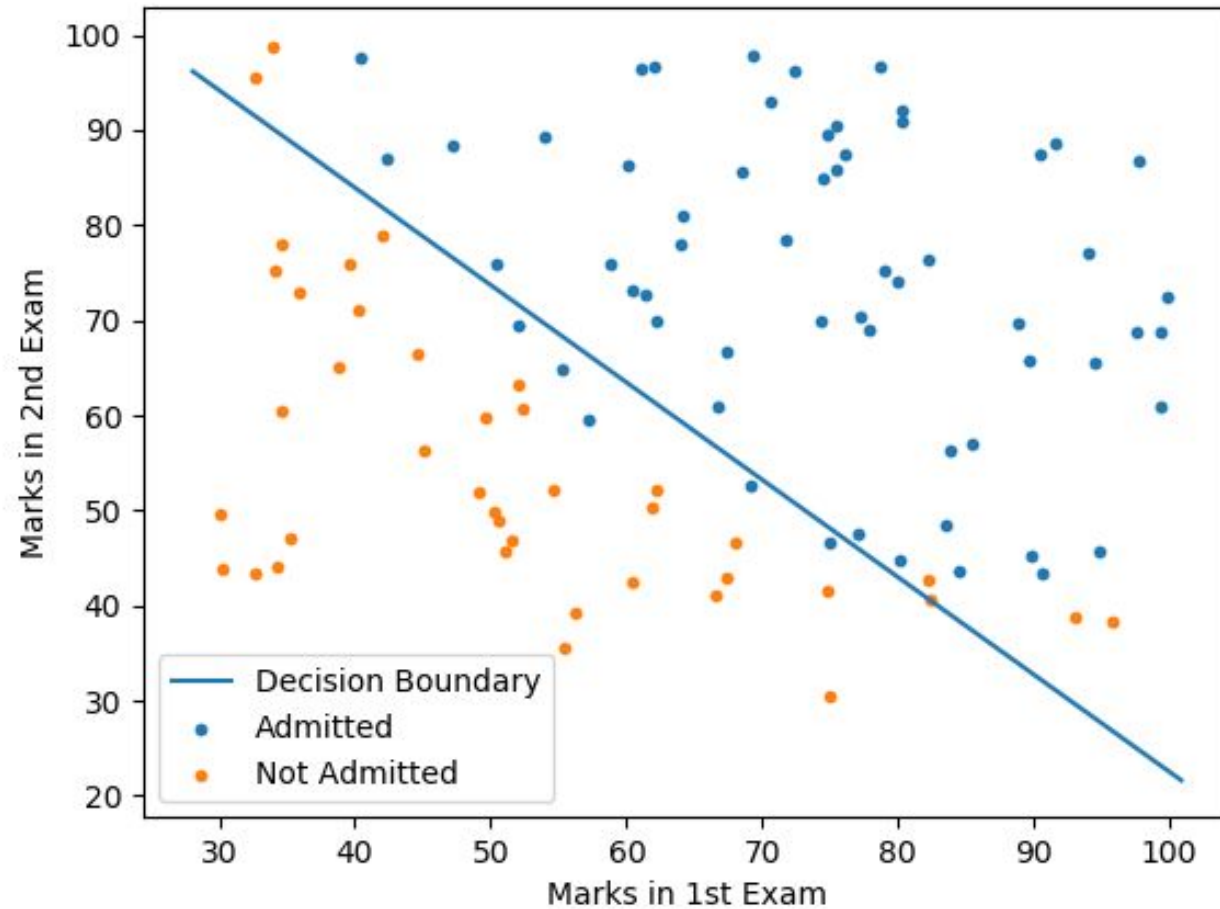
$$h(x) = \frac{1}{1 + e^{-\theta^T x}}$$

$$h(x) = \begin{cases} > 0.5, & \text{if } \theta^T x > 0 \\ < 0.5, & \text{if } \theta^T x < 0 \end{cases}$$



$$\text{cost} = \begin{cases} -\log(h(x)), & \text{if } y = 1 \\ -\log(1 - h(x)), & \text{if } y = 0 \end{cases}$$

$$h(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$
$$x_2 = -\frac{\theta_0 + \theta_1 x_1}{\theta_2}$$



Преимущества:

Благодаря своей эффективности и простоте

- не требует высокой вычислительной мощности
- легко реализуется
- легко интерпретируется
- широко используется аналитиками данных и учеными
- не требует масштабирования функций
- логистическая регрессия дает балл вероятности для наблюдений.

Недостатки:

- не готова обрабатывать избыточное количество категорических признаков/переменных
- подвержена переобучению
- не может решить нелинейную проблему с логистической регрессией, поэтому она требует изменения нелинейных признаков
- не будет хорошо работать с независимыми переменными, которые не коррелируют с целевой переменной и очень похожи или коррелированы друг с другом.



ГУАП

guap.ru

Спасибо за внимание!

Авторы: Вишневский Яков, Герасимец Анна