

Машинное обучение в экономике

Нейронные сети

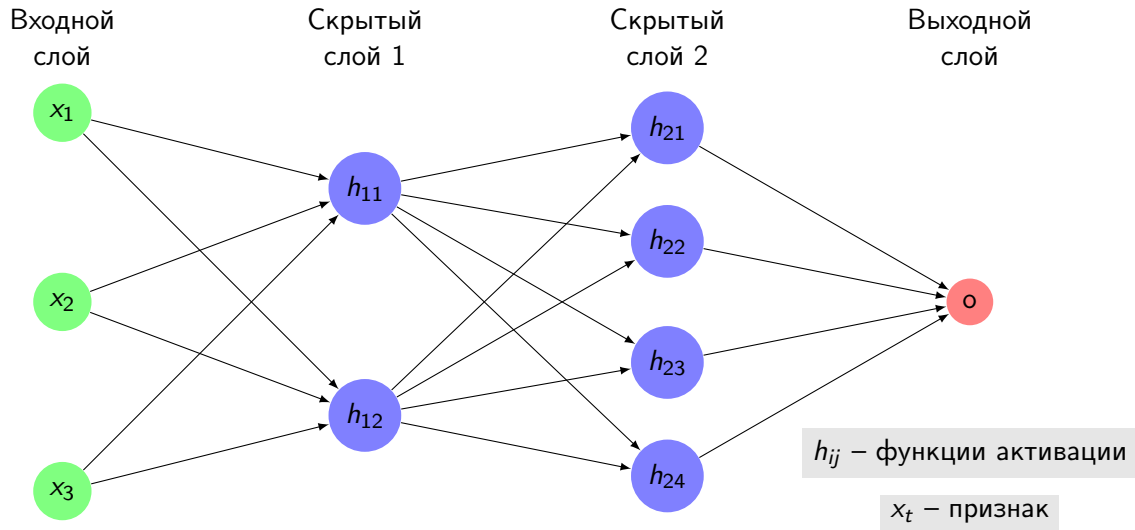
Потанин Богдан Станиславович

доцент, научный сотрудник, кандидат экономических наук

2023–2024

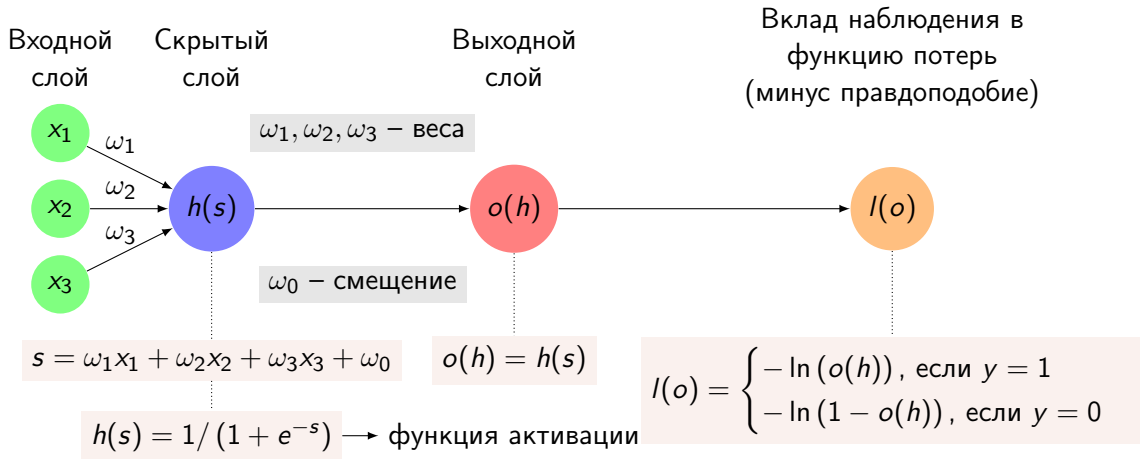
Нейронная сеть

Графическая репрезентация идеи



Нейронная сеть

Логистическая регрессия как нейронная сеть



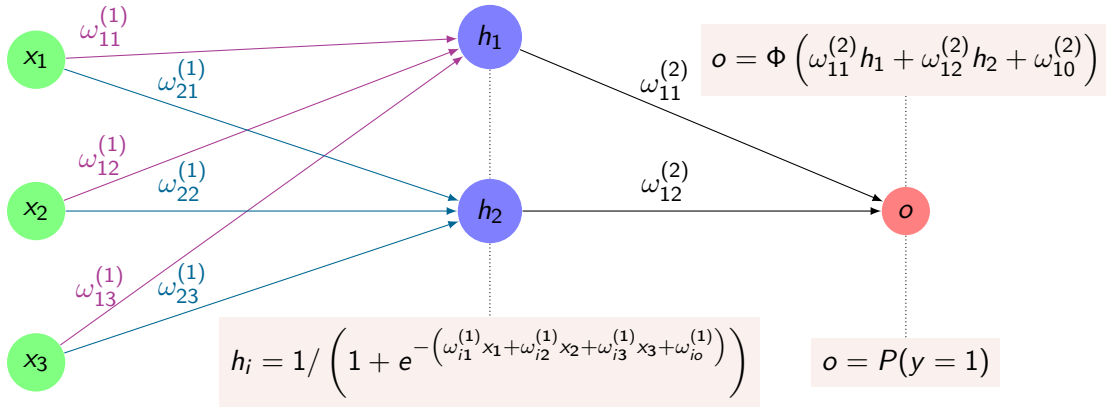
Нейронная сеть

Пример с двумя функциями активации

Входной слой

Скрытый слой

Выходной слой



- Определяется функция потерь. Например, в задаче классификации в качестве функции потерь может выступать умноженная на минус единицу функция правдоподобия. При работе с непрерывными переменными можно рассмотреть сумму квадратов отклонений предсказанных значений от истинных.
- Сперва представим, что веса зафиксированы. В таком случае для получения значения функции потерь достаточно каждое наблюдение провести через нейросеть (от входного слоя до выходного слоя) и воспользоваться полученными значениями (из выходного слоя) для расчета функции потерь. В задаче классификации эти значения, как правило, представляют собой вероятности, а при работе с непрерывными переменными – предсказанные значения целевой переменной.
- Можно воспользоваться любым удобным алгоритмом численной оптимизации, для того чтобы найти веса $\omega_{(ij)}^k$ и смещения $\omega_{(i0)}^k$ нейросети, минимизирующие функцию потерь.
- Число слоев и нейронов, а также функции активации являются гиперпараметрами нейросети.

Функции активации

Общая идея

- Как правило во всех скрытых слоях применяется одна и та же функция активации. Однако, можно использовать разные функции активации даже в рамках одного слоя, а также включать дополнительные оцениваемые параметры этих функций активации.
- Наиболее популярные функции активации:

Сигмоида (логистическая): $1 / (1 + e^{-s})$

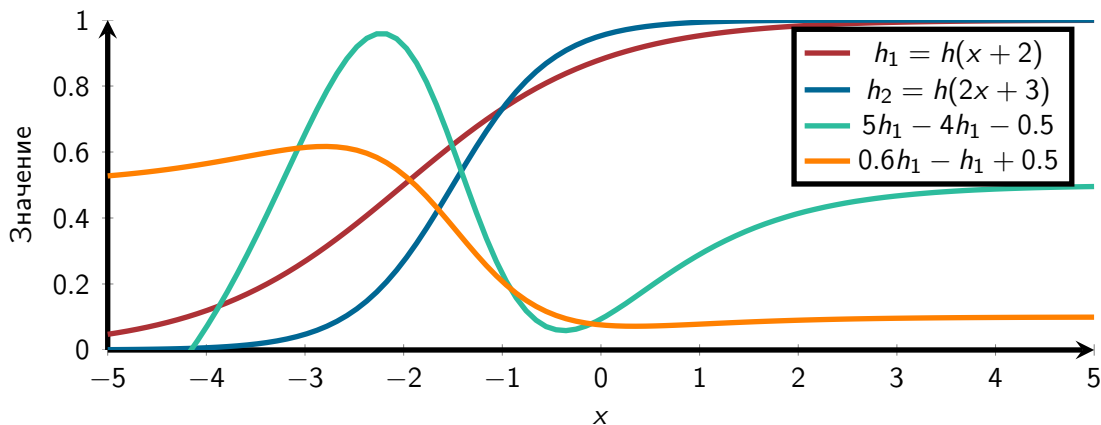
ReLU: $\max(0, s)$

ELU:
$$\begin{cases} \alpha(e^s - 1), & \text{если } s \leq 0 \\ s, & \text{если } s > 0 \end{cases}$$

- Форма каждой функции активации варьируется достаточно слабо, однако линейные комбинации даже одних и тех же функций могут принимать самые разнообразные формы, благодаря чему и достигается гибкость нейросети.

Функции активации

Графический пример с одним признаком и линейной комбинацией логистических функций активации



Вывод – h_1 и h_2 похожи, но их линейные комбинации могут принимать разнообразные формы, что позволяет аппроксимировать очень сложные зависимости.