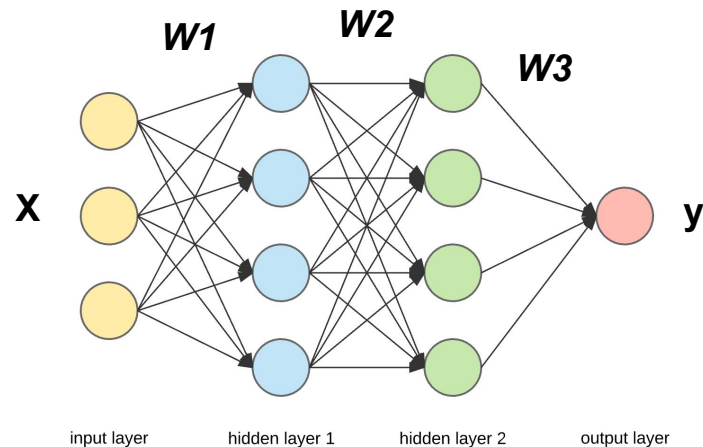


Многоклассовая классификация

Основы Deep Learning

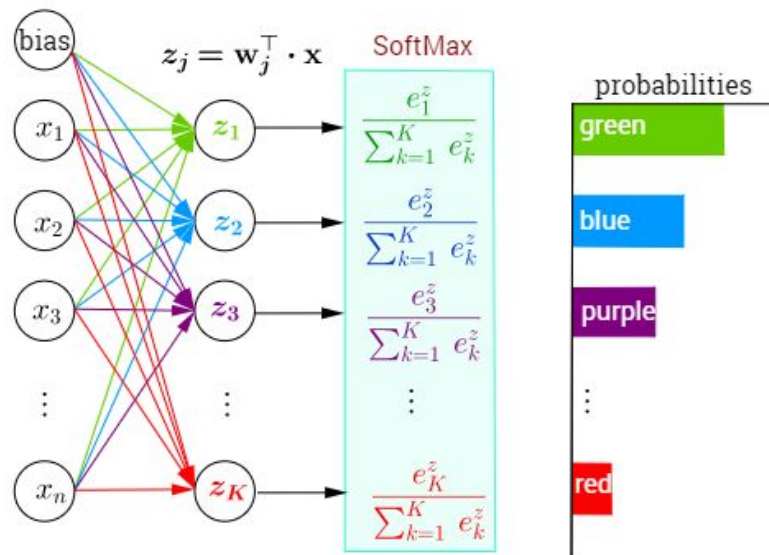
Полносвязная нейросеть

- Вход:
Числовая матрица X размера (n, k)
- Внутри:
Матрицы параметров (веса нейросети)
- Выход: вектор ответов y размера $(n, 1)$
 - метки классов (классификация)
 - вещественные числа (регрессия)



Многоклассовая классификация

- В случае **K** классов: **K** нейронов на выходном слое, поскольку хотим конечный размер матрицы $A[\text{out}]$ равным (n, K)
- Выходной слой не имеет активаций
- Строка матрицы $A[\text{out}]$ с номером i -- это K чисел, так называемы score 'ы классов



Многоклассовая классификация

- Применяем к выходной матрице Softmax, чтобы score'ы перешли в вероятности классов
- Теперь число с номером j в i -ой строке -- это вероятность принадлежности объекта i к классу j (сумма в каждой строке равна единице)

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{k=1}^K e^{z_k}}$$

$z[i]$ -- i -ое значение в строке матрицы $A[\text{out}]$
(и так для всех строк по отдельности)

Многоклассовая классификация

- Применяем к выходной матрице Softmax, чтобы score'ы перешли в вероятности классов
- Теперь число с номером j в i -ой строке – это вероятность принадлежности объекта i к классу j (сумма в каждой строке равна единице)



Пример преобразования строки score'ов одного объекта
(здесь строка вытянута в столбец)

Многоклассовая классификация

- Применяем к выходной матрице Softmax
- Чтобы score'ы перешли в вероятности классов
- Теперь число с номером j в i -ой строке -- это вероятность принадлежности объекта i к классу j
- Результат подаётся в CrossEntropy лосс, где \mathbf{p} -- матрица предсказаний, \mathbf{t} -- истинные метки

$$\alpha_{ij} = \text{softmax}_j(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k \in \mathcal{N}_i} \exp(e_{ik})}.$$

$$\text{cross-entropy} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^k t_{i,j} \log(p_{i,j})$$