Задача 1. 1. Рассмотрим простейший перцептрон с константой, одним входом x_1 и пороговой функцией активации. Подберите веса так, чтобы перептрон реализовывал логическое отрицание (в ответ на 0 выдавал 1 и наоборот).

2. Рассмотрим простейший перцептрон с константой, двумя входами x_1 и x_2 и пороговой функцией активации. Подберите веса так, чтобы перцептрон реализовывал логическое ИЛИ.

3. докажите, что невозможно подобрать веса так, чтобы перцептрон реализовывал исключающее логическое ИЛИ (XOR).

4. придумайте, какой признак можно добавить на вход перцептрону, чтобы можно было реализовать XOR. Подберите веса.

Задача 2. Возьмём несколько перцептронов с одним входом и применим к их выходам ещё один перцептрон.

1. Пусть у нас есть три объекта x_i. Покажите, что можно подобрать веса перцептронов так, чтобы восстановить любой набор меток y_i. Сколько перцептронов в первом слое нужно?

2. Можно ли с помощью этой же архитектуры восстановить любой набор меток на 4 объектах?

Задача 3. Пусть в каждой точке $x \in X$ пространства объектов задана вероятность $p(y = +1|x)$ того, что данный объект относится к классу $+1$, и пусть алгоритм $b(x)$ возвращает числа из отрезка $[0, 1]$. Потребуем, чтобы эти предсказания пытались в каждой точке x приблизить вероятность положительного класса $p(y = +1|x)$.

Это требование можно записать через минимизацию матожидания функции потерь $L(y, b(x))$:

$$\min_b \mathbb{E}_y[L(y, b(x))|x].$$

1. Покажите, что квадратичная функция потерь $L(y, b) = ([y = +1] - b)^2$ позволяет предсказывать корректные вероятности.

2. Покажите, что функция потерь $L(y, b) = |[y = +1] - b|$ не позволяет предсказывать корректные вероятности.

Задача 4. Рассмотрим метрику “доля дефектных пар”. Пусть классификатор возвращает для каждого объекта выборки X оценку $b(x_i) \in [0, 1]$. Отсортируем объекты по возрастанию этих оценок $b(x_{(1)}) < \ldots < b(x_{(n)})$.Пусть y_i — истинный ответ на объекте x_i. Тогда метрика определяется следующим образом:

$$\mathrm{DP}(b, X) = \frac{2}{n(n-1)} \sum_{i<j} [y_{(i)} > y_{(j)}]$$

Как эта метрика связана с ROC-AUC?

Задача 5. Рассмотрим целевую функцию логистической регрессии с константой

$$Q(w) = \frac{1}{l}\sum L(y_i, b_i),$$

где $b_i = (1 + \exp(-\langle w, x_i\rangle))^{-1}$ и $L(y_i, b_i) = \begin{cases} -\log(b_i) & y_i = 1 \\ -\log(1-b_i) & y_i = 0 \end{cases}$.

1. Найдите $dQ(w)$ и $d^2Q(w)$;
2. Найдите $dQ(0)$ и $d^2Q(0)$;
3. Выпишите квадратичную аппроксимацию для $Q(w)$ в окрестности $w = 0$;
4. С какой задачей совпадает задача минимизации квадратичной аппроксимации?

Задача 6. Винни-Пух знает, что мёд бывает правильный ($honey_i = 1$), и неправильный ($honey_i = 0$). Пчёлы также бывают правильные ($bee_i = 1$), и неправильные ($bee_i = 0$). По 100 своим попыткам добыть мёд Винни-Пух составил таблицу сопряженности:

	$bee_i = 1$	$bee_i = 0$
$honey_i = 1$	12	36
$honey_i = 0$	32	20

Винни-Пух использует логистическую регрессию с константой для прогнозирования правильности мёда с помощью правильности пчёл.

1. Какие оценки коэффициентов получит Винни-Пух?
2. Какой прогноз вероятности правильности мёда при встрече с неправильными пчёлами даёт логистическая модель? Как это число можно посчитать без рассчитывания коэффициентов?