***Логистическая регрессия.***

Основная идея логистической регрессии заключается в том, что пространство исходных значений может быть разделено линейной границей (т. е. прямой) на две соответствующих классам области.

Уравнение

$$\sum\_{i=1}^{n}w\_{i}x\_{i}=u$$

задаёт (n-1)-гиперплоскость в n-мерном пространстве. Пространство разбивается на две области

$$\sum\_{i=1}^{n}w\_{i}x\_{i}>u и \sum\_{i=1}^{n}w\_{i}x\_{i}<u$$

соответствующие разделению множества объектов на два класса (линейный классификатор).

 Этот классификатор можно реализовать через один нейрон. Ограничимся двумерным случаем:

$$I\left(w\_{0}x+w\_{1}y+b>0\right)$$

Значение 0 или 1 на выходе соответствуют принадлежности к одному из двух кластеров.

***Относительно значений весов*** $w\_{0} w\_{1}$ ***и смещения bias:***

 Для простоты – два кластера с одинаковым распределением.

 Если известны центры кластеров $\left(X\_{1},Y\_{1}\right) и \left(X\_{2},Y\_{2}\right)$, то прямая проходящая через эти центры имеет вид:

$$\frac{\left(x-X\_{1}\right)}{(X\_{2}-X\_{1})}=\frac{\left(y-Y\_{1}\right)}{(Y\_{2}-Y\_{1})} или y=ax+b, где a=\frac{\left(Y\_{2}-Y\_{1}\right)}{(X\_{2}-X\_{1})}$$

$$y=\frac{\left(x-X\_{1}\right)(Y\_{2}-Y\_{1})}{(X\_{2}-X\_{1})}+Y\_{1}=\left(y-Y\_{1}\right)$$

Координаты центр соответствующего отрезка равны

$$X\_{0}=\frac{X\_{1}+X\_{2}}{2} Y\_{0}=\frac{Y\_{1}+Y\_{2}}{2}$$

Уравнение прямой, проходящей через точку $\left(X\_{0},Y\_{0}\right)$ и перпендикулярной прямой проходящей через центры кластеров имеет вид:

$$\frac{\left(y-Y\_{0}\right)}{\left(x-X\_{0}\right)}=-\frac{1}{a} или x+ay-X\_{0}-aY\_{0}=0 $$

Этому уранению следует придать симметричный относительно x-y вид, в этом случае мы избежим неприятностей при $X\_{2}=X\_{1}$.

$$\left(X\_{2}-X\_{1}\right)x+\left(Y\_{2}-Y\_{1}\right)y-\left(X\_{2}-X\_{1}\right)X\_{0}-\left(Y\_{2}-Y\_{1}\right)Y\_{0}=0$$



 Если кластеры одинаковы по размеру, то эта прямая будет являться разделителем в линейном классификаторе. В этом случае для весов $w\_{0} w\_{1}$ и смещения *b* получаем следующие значения:

$$w\_{0}=X\_{2}-X\_{1}, w\_{1}=Y\_{2}-Y\_{1}, bias=-\left(X\_{2}-X\_{1}\right)X\_{0}-\left(Y\_{2}-Y\_{1}\right)Y\_{0}$$

***Обучение нейросети – метод Монте-Карло***

Подбираем значения $w\_{0},w\_{1},bias$ из условия наилучшего совпадения предсказания сети с реальными значениями. Из уравнений

$$\left\{\begin{array}{c}X\_{2}-X\_{1}=w\_{0}\\Y\_{2}-Y\_{1}=w\_{1}\\-\left(X\_{2}-X\_{1}\right)X\_{0}-\left(Y\_{2}-Y\_{1}\right)Y\_{0}=bias\end{array}\right.$$

находим

$$a=\frac{\left(Y\_{2}-Y\_{1}\right)}{(X\_{2}-X\_{1})}=\frac{w\_{1}}{w\_{0}}$$

$$Y\_{0}=-\frac{1}{a}X\_{0}-\frac{bias}{w\_{1}}$$



***Обучение нейросети – функция ошибок, целевая функция***

Входные данные: $x\_{i} 0\leq i<N$

Выходные данные: $t\_{i} 0\leq i<N$

Выход сети: $y\_{i} 0\leq i<N$

Задача: подобрать параметры сети так, чтобы минимизировать значение

$$F\left(w, b\right)=\sum\_{i}^{}\left(y\_{i}-t\_{i}\right)^{2}$$