One of Pioneer of Neutral Network：Yann LeCun

Multi-class classification:

Suppose we have four class;

The output layer is the essentially four logistic regression classifiers. Each of them will capture one of the class.

Training set is the (x1,y1), (x2,y2),(x3,y3)….(xm,ym)

Y is [1,0,0,0]/[0,1,0,0]/[0,0,1,0]/[0,0,0,1] to represent the four classes.