



Shallow neural network: $f(x,y) = \phi(a_1x + a_2y - b)$

Score: $J(a_1, a_2, b) = \sum_{i=1}^n z_i \phi(a_1x_i + a_2y_i - b) - c(a_1^2 + a_2^2)$

then $\max_{a_1, a_2, b} J(a_1, a_2, b)$

\hookrightarrow penalizes large $\|(a_1, a_2)\| = \sqrt{a_1^2 + a_2^2}$

Some $c > 0$
e.g. $c = 1$