original representation   intermediate features

$z_1 = \phi_1(x)$

$z_2 = \phi_2(x)$

$z_S = \phi_S(x)$

x_1
x_2
x_d

Classifier $\longrightarrow$ **Ragù**

x_1
x_2
x_d

$$\sum_j W_{ij} x_j + b_i$$

$\phi_i(x)$
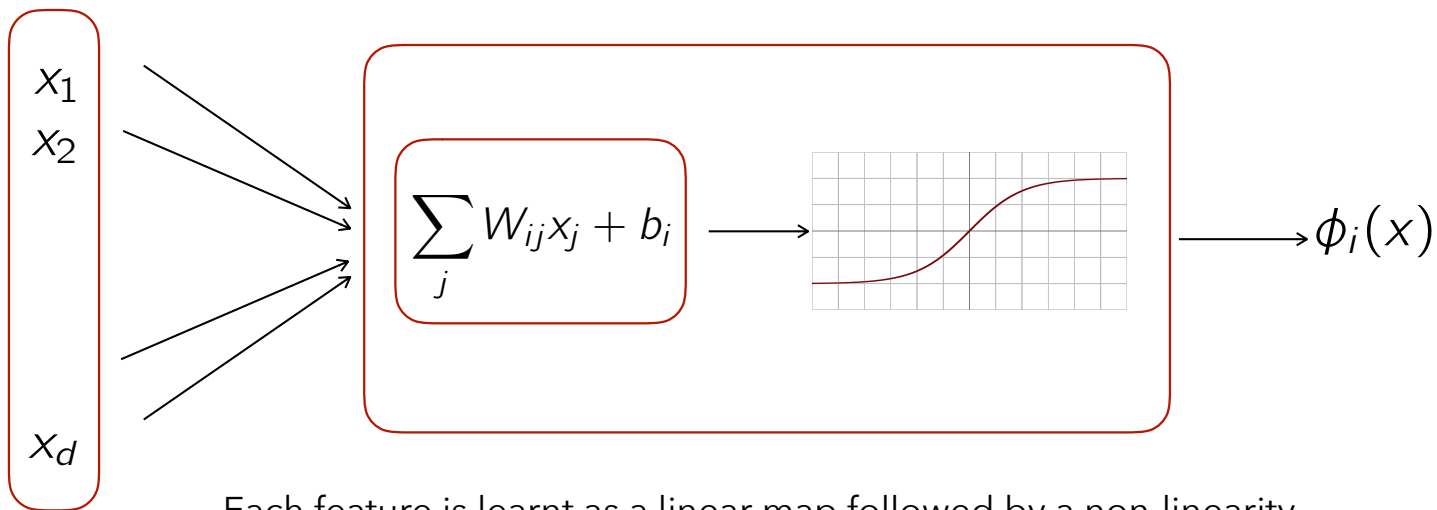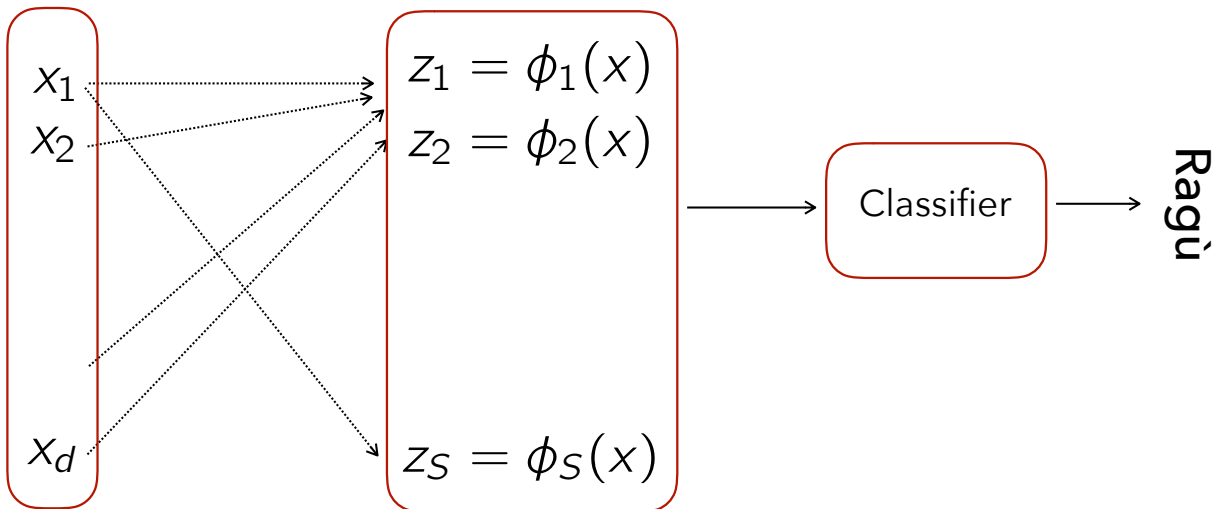
Each feature is learnt as a linear map followed by a non-linearity