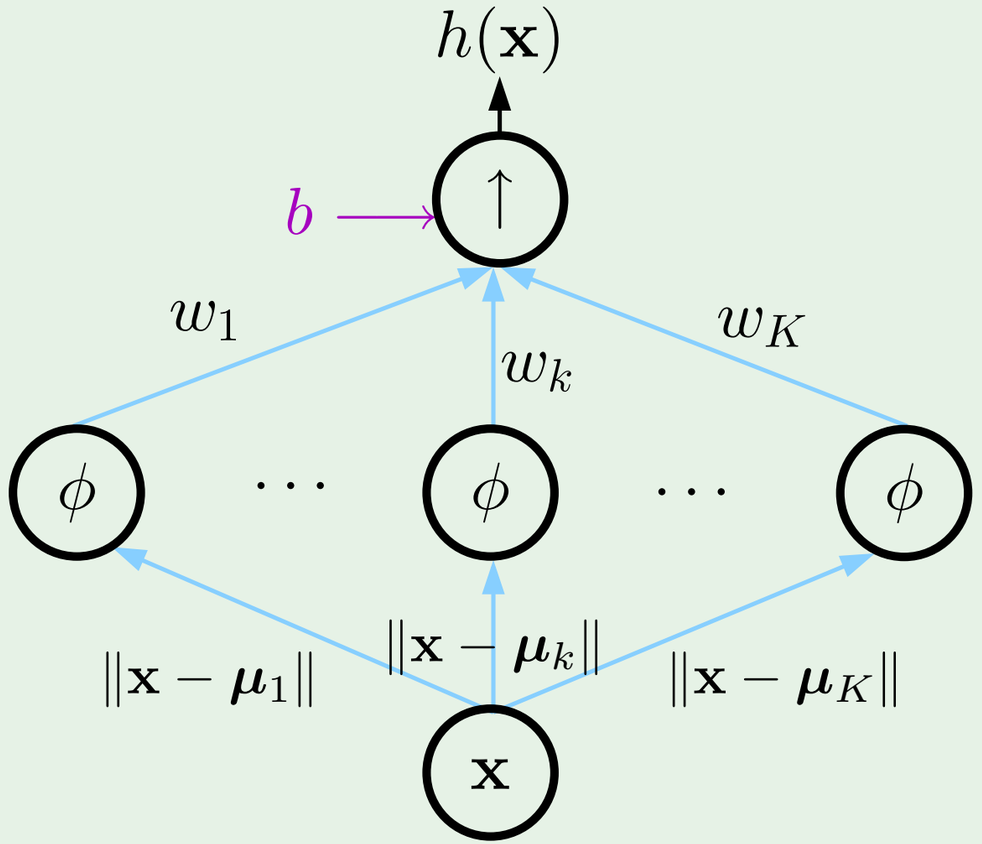
RBF network

The "features" are $\exp\left(-\gamma \left\|\mathbf{x} - \boldsymbol{\mu}_k\right\|^2\right)$

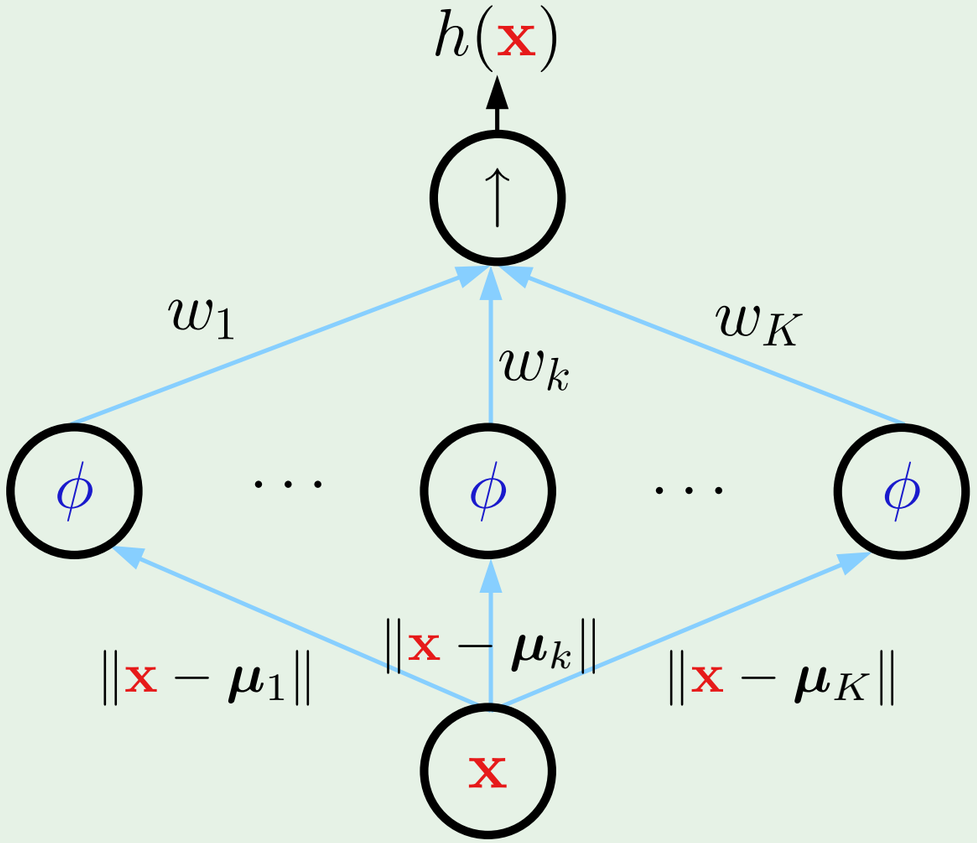
Nonlinear transform depends on $\mathcal{D}$

$\Longrightarrow$ No longer a linear model

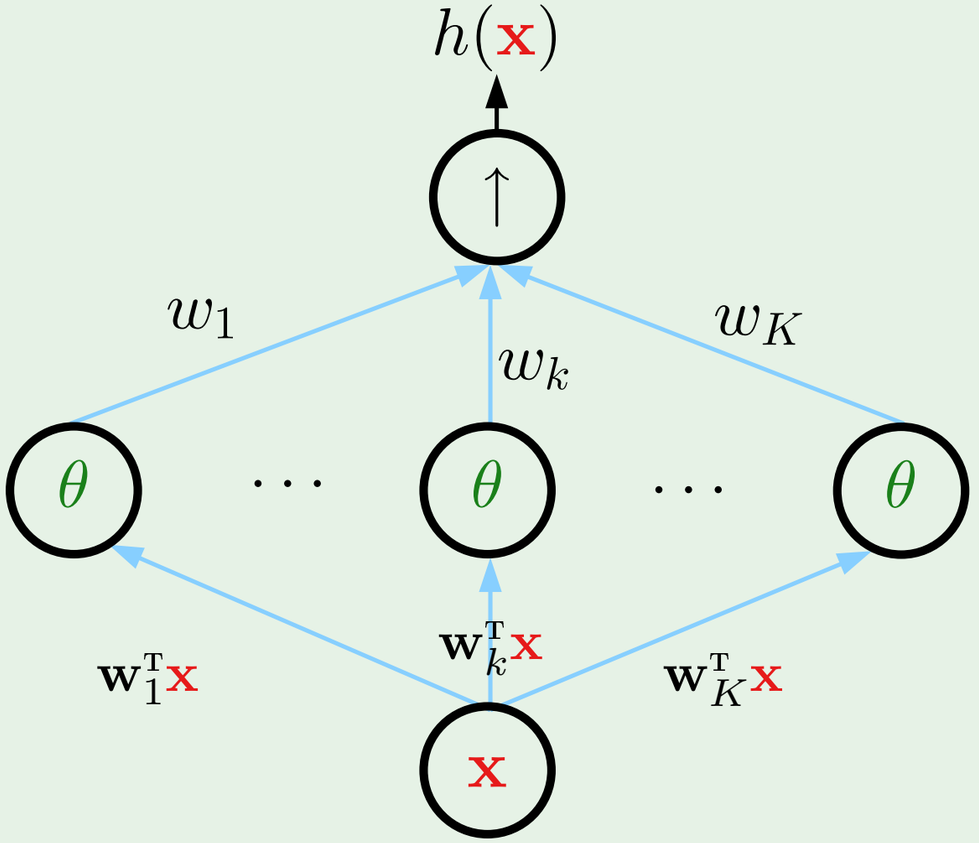


A bias term (b or w_0) is often added

Compare to neural networks



RBF network



neural network

Choosing γ

Treating γ as a parameter to be learned $\quad h(\mathbf{x}) = \sum_{k=1}^{K} w_k \exp\left(-\gamma \|\mathbf{x} - \boldsymbol{\mu}_k\|^2\right)$

Iterative approach ($\sim$ **EM algorithm** in mixture of Gaussians):

1. Fix γ, solve for $w_1, \cdots, w_K$
2. Fix $w_1, \cdots, w_K$, minimize error w.r.t. γ

We can have a different γ_k for each center $\boldsymbol{\mu}_k$