



Why

A distribution/density selector picks a probabilistic model for a dataset by using the maximum likelihood principle. We want to generalize this idea.¹

Definition

An *unsupervised data model* is a function $\ell : \mathbf{R}^d \rightarrow \mathbf{R}$ where $\ell(x)$ is the *surprise* of the vector x .²

¹Future editions probably include a genetic approach, which moves through clustering and mixtures of Gaussians to illustrate the point more fully.

²Future editions will expand.

