maschine learning Algorithm learns class of tasks, measured by loss function, from experience.

supervised learning learn $h : \Delta^* \to \Sigma^*, h = t$; example: $(x, y) \in \Delta^* \times \Sigma^*,\ t(x) = y$.

unsupervised learning learn $h : \Delta^* \to \Sigma^*, \ker(h) = \ker(t)$; example: $x \in \Delta^*$.

reinforcement learning learn strategy based on feedback from environment.

1 Supervised Learning

- model function $t : \mathcal{M} \to \mathcal{R}$
- $supp(t) = \{m \in \mathcal{M} \mid t(m) \neq 0\}$

Hypothesis of A: potential result of A

Hypothesis space $\mathcal{H}_A$ of A: set of all hypotheses

h fits D if $h(x_i) = y_i$ for all $(x_i, y_i) \in D$

Version space $\mathcal{V}_A(D)$ of A: all hypotheses that fit D

Inductive bias of A: set of assumptions that A uses to predict outputs of unseen data

Conjunctive Clause $\theta = (\theta_1, ..., \theta_k), \theta_i \in M_i \cup \{\star, \bot\}$

- $\theta_\bot = (\bot, ..., \bot)$ most specific
- $\theta_\star = (\star, ..., \star)$ most general
- $supp(h_{\theta_\bot}) = \emptyset, supp(h_{\theta_\star}) = \mathcal{M}$
- $h_{\theta_\bot} = h_{(\theta_1,...,\bot,...,\theta_k)=...}$

induced hypothesis $h_\theta(m_1, ..., m_k) = 1$ if $\forall i : \theta_i \in \{m_i, \star\}$ else 0

loss functions (and derivatives)

- $l(h, D) = \sum_{i=1}^{n}(1 - \delta_{y_i, h(x_i)})$
- $\delta_{ij} = 1$ if $i = j, 0$ otherwise.
- asd