Hasso Plattner Institut
Digital Engineering • Universität Potsdam

Computational Learning Theory
1.3 Hypothesis Space

Karen Seidel
Hasso-Plattner-Institute
University of Potsdam

Homogeneous Outputformat

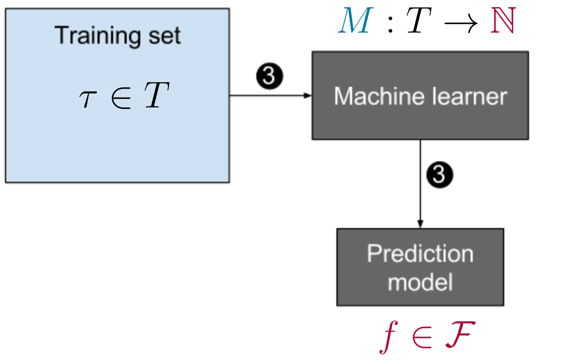


Image from: `https://commons.wikimedia.org/wiki/File:Machine_learning_nutshell_--_Train_a_machine_learning_model.svg`

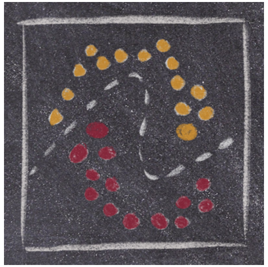
Definition (Hypothesis Space, Learner)

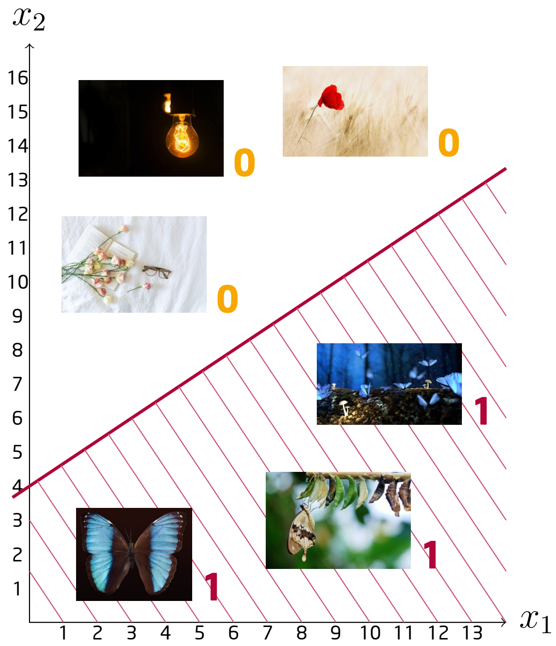
A set $\mathcal{F} = \{f_i \mid i \in \mathbb{N}\}$ of prediction models is called *hypothesis space* if there is a computer program which on an array $[i, n]$ of non-negative integers returns 1 if $f_i(n) = 1$ and 0 otherwise.

A *machine learner* M is a computable function $M : T \to \mathbb{N}$ with the output interpreted with respect to a prefixed hypothesis space $\mathcal{F}$.

Hypothesis Space of SVMs and Perceptron

- Input: training data $((\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_t, y_t))$.
- Output: halfspace.
- Realizability Assumption: data linearly separable. (No-Free-Lunch-Theorem)
- Many learning algorithms use halfspaces as prediction models, sometimes combined with a kernel function.

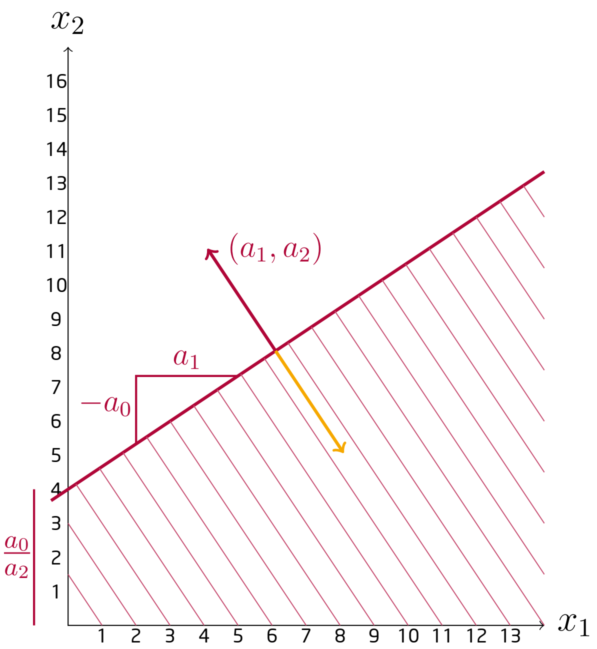




The Hypothesis Space of Halfspaces I

- The line is given by the formula $x_2 = \frac{2}{3}x_1 + 4$ or equivalently $12 = -2x_1 + 3x_2$.
- The normal vector $(-2, 3)$ points outside the halfspace.
- The line $-12 = 2x_1 - 3x_2$ decribes the halfspace on the other side of the same line.

We consider the set of all halfspaces given by a line $a_0 = a_1x_1 + a_2x_2$ with $a_0, a_1, a_2 \in \mathbb{Z}$.



The Hypothesis Space of Halfspaces II

We consider the set of all halfspaces given by a line $a_0 = a_1x_1 + a_2x_2$ with $a_0, a_1, a_2 \in \mathbb{Z}$.

- Given $a_0 = 12, a_1 = -2, a_2 = 3$ and vector $(x_1, x_2) = (-2, 7)$, a computer can decide whether the vector lies in the halfspace $L_{12,-2,3} = \{(x_1, x_2) \mid 12 \geq -2x_1 + 3x_2\}$ given by the line $12 = -2x_1 + 3x_2$.
- It follows the same method for other values of a_0, a_1, a_2 and x_1, x_2.
- The decision procedure is in this sense universal in the inputs (a_0, a_1, a_2) and (x_1, x_2).

How is a halfspace represented by a natural number?

The Hypothesis Space of Halfspaces III

We consider the set of all halfspaces given by a line $a_0 = a_1x_1 + a_2x_2$ with $a_0, a_1, a_2 \in \mathbb{Z}$.

How is a halfspace represented by a natural number?

- The computer can identify each integer with a natural number

$\mathbb{Z}$	0	-1	1	-2	2	-3	3	-4	4	$\ldots$
$\mathbb{N}$	0	1	2	3	4	5	6	7	8	$\ldots$

- Applying the merge method, the computer can for inputs $i, n \in \mathbb{N}$ reconstruct (i_0, i_1, i_2) and (n_1, n_2) corresponding to (a_0, a_1, a_2) and (x_1, x_2), respectively.
- Then it does run the uniform decision procedure to decide whether $(x_1, x_2) \in L_{a_0,a_1,a_2}$.
- The computer does return 1 if the answer is yes and 0 otherwise. This defines $f_i(n)$.

Homogeneous Outputformat

Definition (Hypothesis Space, Learner)

A set $\mathcal{F} = \{f_i \mid i \in \mathbb{N}\}$ of prediction models is called *hypothesis space* if there is a computer program which on an array $[i, n]$ of non-negative integers returns 1 if $f_i(n) = 1$ and 0 otherwise.

A *machine learner* M is a computable function $M : T \to \mathbb{N}$ with the output interpreted with respect to a prefixed hypothesis space $\mathcal{F}$.

The set of halfspaces is a hypothesis space.

- The formal definitions cover the binary classification of images with halfspaces.
- There are a lot of other hypothesis spaces, e.g., the set of weight vectors for a fixed neural network architecture.

Combinatorial Analysis of Learning

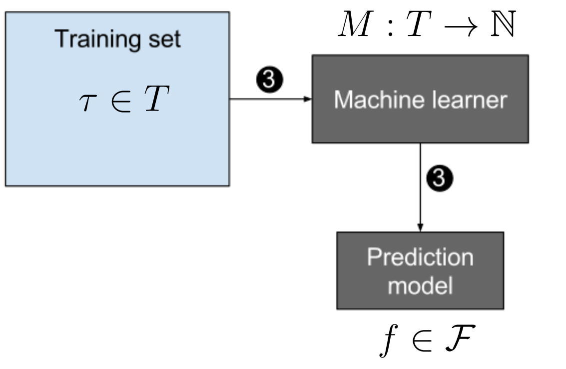


Image from: `https://commons.wikimedia.org/wiki/File:Machine_learning_nutshell_--_Train_a_machine_learning_model.svg`

Is there an M that *learns* a well performing $f \in \mathcal{F}$
from sequences $\tau \in T$ of increasing length?