Machine Learning Exercise 01 Solution

Eugen Sawin

May 10, 2012

Exercise 1.1

(a) Possible Pacman tasks that could be solved using machine learning are:

- Survival: learn to escape the ghosts.
- Fight back: learn to eat ghosts when they are vulnerable.
- Eat: maximise dots eaten per elapsed time.

(b) I would apply supervised learning on:

- Predict tomorrows price of a particular stock
- Predict your life expectancy
- Predict if it is going to rain tomorrow

(c) The following problems are classification problems:

- Predict if it is going to rain tomorrow

Exercise 1.2

(a) A version space consists of all consistent hypothesis from a hypothesis space with respect to the given set of examples. The order is induced by the strict ordering *more-general-than*($>_g$)/*more-specific-than*($<_g$).

S represents the set of the most specific consistent hypothesis (summary of all positive examples) and G represents the set of the most general consistent hypothesis (summary of all negative examples). All other consistent hypothesis are between these given boundaries, implicitely.

For inconsistent examples S and G become empty and so does the version space, this is also the case if the language used for the definition of the hypothesis space is insufficient to describe the target concept.

(b) Just imagine a beautiful looking graph here depicting the ordered hypothesis space.

(c)

i	S_i	G_i
0	$\{\langle\emptyset,\emptyset,\emptyset\rangle\}$	$\{\langle ?,?,?\rangle\}$
1	$\{\langle +,+,+\rangle\}$	$\{\langle ?,?,?\rangle\}$
2	$\{\langle +,+,?\rangle\}$	$\{\langle ?,?,?\rangle\}$
3	$\{\langle +,+,?\rangle\}$	$\{\langle ?,+,?\rangle\}$
4	$\{\langle +,+,?\rangle\}$	$\{\langle +,+,?\rangle\}$
5	$\{\langle +,+,?\rangle\}$	$\{\langle +,+,?\rangle\}$
6	$\{\langle +,+,?\rangle\}$	$\{\langle +,+,?\rangle\}$

(d) No, but it does affect the convergence speed. The hypothesis space represents a systematic model of all consistent hypothesis and does not converge to local specialisations.

(e) For concrete target concept like $\langle +,+\rangle$ three examples are enough, one positive and two negative. For a

concrete general concept like $\langle +, ? \rangle$ we need four examples, two of each type. For the trivial concept $\langle ?, ? \rangle$ two (obviously) positive examples are enough.

(f)

i	S_i	G_i
0	$\{\langle \emptyset, \emptyset, \emptyset . \emptyset \rangle\}$	$\{\langle ?, ?, ?, ? \rangle\}$
1	$\{\langle +, +, +, - \rangle\}$	$\{\langle ?, ?, ?, ? \rangle\}$
2	$\{\langle +, +, ?, - \rangle\}$	$\{\langle ?, ?, ?, ? \rangle\}$
3	$\{\langle ?, ?, ?, ? \rangle\}$	$\{\langle ?, ?, ?, ? \rangle\}$

It converged to a wrong hypothesis. If we move example d_3, it will result in an empty set S and therefore not consistent hypothesis, because the concept to be learned is not in the defined hypothesis space.

Exercise 1.3

(a) $E(S) = -\frac{1}{7}\log_2\frac{1}{7} - \frac{2}{7}\log_2\frac{2}{7} - \frac{2}{7}\log_2\frac{2}{7} - \frac{2}{7}\log_2\frac{2}{7} \approx 1.95$
$E(S|fever) = \frac{2}{7}(1) + \frac{3}{7}(2) + \frac{2}{7}(1) \approx 1.64$
$E(S|vomiting) = \frac{4}{7}(-2\frac{1}{4}\log_2\frac{1}{4} - \frac{2}{4}\log_2\frac{2}{4}) + \frac{3}{7}(-\frac{1}{3}\log_2\frac{1}{3} - \frac{2}{3}\log_2\frac{2}{3}) \approx 1.25$
$E(S|diarrhea) = \frac{3}{7}(-\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3}) + \frac{4}{7}(-\log_2\frac{2}{4}) \approx 0.96$
$E(S|shivering) = \frac{6}{7}(-2\frac{1}{6}\log_2\frac{1}{6} - 2\frac{2}{6}\log_2\frac{2}{6}) + \frac{1}{7}\cdot 0 \approx 1.64$
As we can clearly see, we get the greates gain with diarrhea $G(S, diarrhea) = 1.95 - 0.96 = 0.99$. Since the human brain is quite good at approximating entropy, this was visible from the beginning and I will refrain from calculating it for each step. So here is the tree.

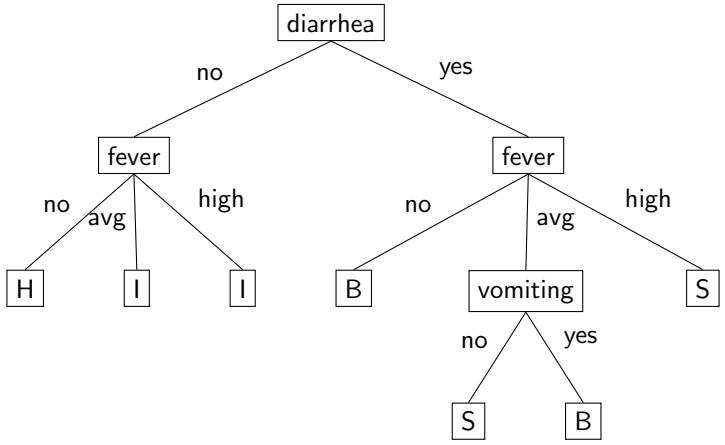

Figure 1: The complete search tree with the dotted edges depicting the solution path.

(b) I'm just assuming what a disjoint definition is, no it does not, since both average and high fever results in influenza in the case of diarrhea.

(c) Candidates are $A = \frac{164+169}{2}$, $B = \frac{175+176}{2}$, $C = \frac{179+180}{2}$, $D = \frac{184+185}{2}$.
$E(A) = \frac{2}{9}\cdot 0 + \frac{7}{9}(-\frac{4}{7}\log_2\frac{4}{7} - \frac{3}{7}\log_2\frac{3}{7}) \approx 0.77$
$E(B) = \frac{4}{9}\cdot 1 + \frac{5}{9}(-\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5}) \approx 0.98$
$E(C) = \frac{6}{9}(-\frac{2}{6}\log_2\frac{2}{6} - \frac{4}{6}\log_2\frac{4}{6}) + \frac{3}{9}(-\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3}) \approx 0.92$
$E(D) = \frac{8}{9}\cdot 1 + \frac{1}{9}\cdot 0 = 1$
Splitting threshold A gives the lowest entropy and therefore greatest gain $G = 0.99 - 0.77 = 0.22$.