2. Exercise Sheet
Introduction to Machine Learning (WiSe 2020/21)

Anton Lammert
Matrikelnummer: ——

Henrik Leisdon
Matrikelnummer: ——

Tony Zöppig
Matrikelnummer: ——

Bauhaus-Universität Weimar

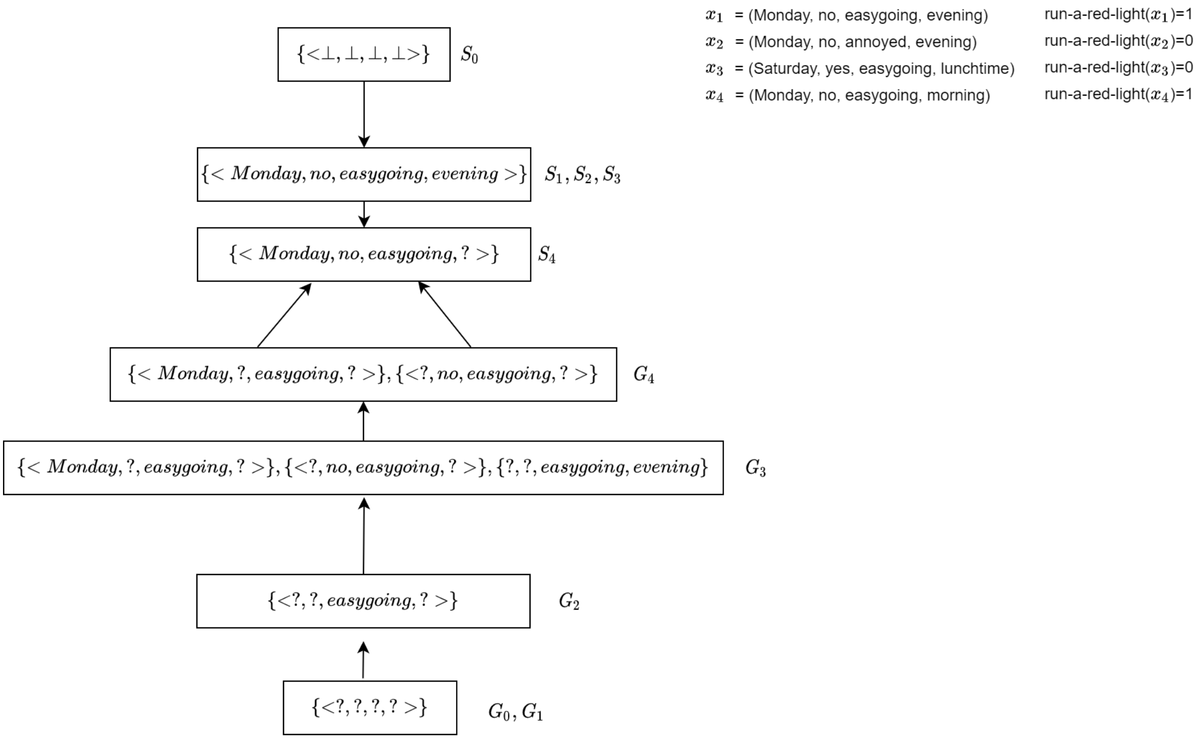
December 4, 2020

Exercise 1

(a) The examples with (x,c(x)=1) in set D are x_1 and x_4. Find-S for training set D is:

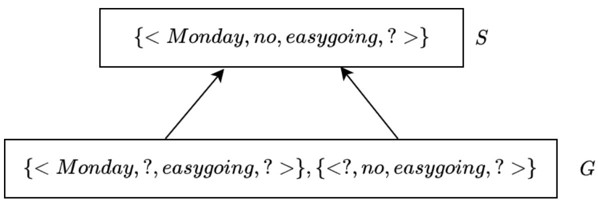
	$h_0 = s_0 = \langle \perp, \perp, \perp, \perp \rangle$
x_1=(Monday, no, easygoing, evening)	$h_1 = \langle$ Monday, no, easygoing, evening $\rangle$
x_2=(Monday, no, annoyed, evening)	$h_2 = \langle$ Monday, no, easygoing, evening $\rangle$
x_3=(Saturday, yes, easygoing, lunchtime)	$h_3 = \langle$ Monday, no, easygoing, evening $\rangle$
x_4=(Monday, no, easygoing, morning)	$h_4 = \langle$ Monday, no, easygoing, ? $\rangle$

(b) Candidate Elimination Algorithm



The boundary set for S is: $\{\langle Monday, no, easygoing, ?\rangle\}$
The boundary set for G is: $\{\langle Monday, ?, easygoing, ?\rangle, \langle ?, no, easygoing, ?\rangle\}$

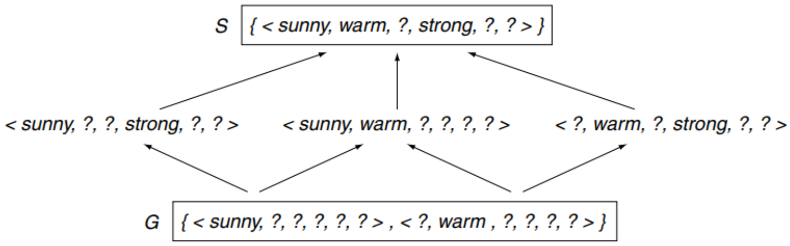
(c) Version Space$_{H,D} = \{h | h \in H \wedge (\forall (x, c(x)) \in D : h(x) = c(x)\}$ so the visual Representation version space for the examples looks like the following:



so it contains the hypothesis: $\{\langle Monday, no, easygoing, ?\rangle, \langle Monday, ?, easygoing, ?\rangle, \langle ?, no, easygoing, ?\rangle\}$.

Exercise 2

(a) Yes, a version space can contain further hypotheses. This is possible if there are consistent hypotheses, which are more general than S and more specific than G. This can be seen in the following picture from the lecture slides:



(b) If s_1, s_2 are two hypothesis from the final set S (the output of the candidate elimination algorithm) then they are specific and no element in S is more general than another element in S.

```
IF c(x) = 1 THEN // x is a positive example.
  FOREACH g ∈ G DO IF g(x) ≠ 1 THEN G = G \ {g} ENDDO
  FOREACH s ∈ S DO
    IF s(x) ≠ 1 THEN
      S = S \ {s}, S+ = min_generalizations(s, x)
      FOREACH s ∈ S+ DO IF (∃g ∈ G : g ≥g s) THEN S = S ∪ {s} ENDDO
      FOREACH s ∈ S DO IF (∃s' ∈ S : s' ≠ s ∧ s ≥g s') THEN S = S \ {s} ENDDO
  ENDDO
ELSE // x is a negative example.
```

- [ ] $(s_2 \geq_g s_1) \vee (s_1 \geq_g s_2)$
- [ ] $(s_2 \geq_g s_1) \wedge (s_1 \geq_g s_2)$
- [x] $(s_2 \not\geq_g s_1) \vee (s_1 \not\geq_g s_2)$
- [x] $(s_2 \not\geq_g s_1) \wedge (s_1 \not\geq_g s_2)$

(c) The Find-S algorithm has the stronger inductive bias. The Candidate Elimination Algorithm classifies only instances that are consistent with all elements in the version space. The Find-S algorithm on the other hand finds the most specific hypothesis consistent with the positive training examples. It then uses this hypothesis to classify all further instances. It is thus more capable of generalization and thus has the higher inductive bias.

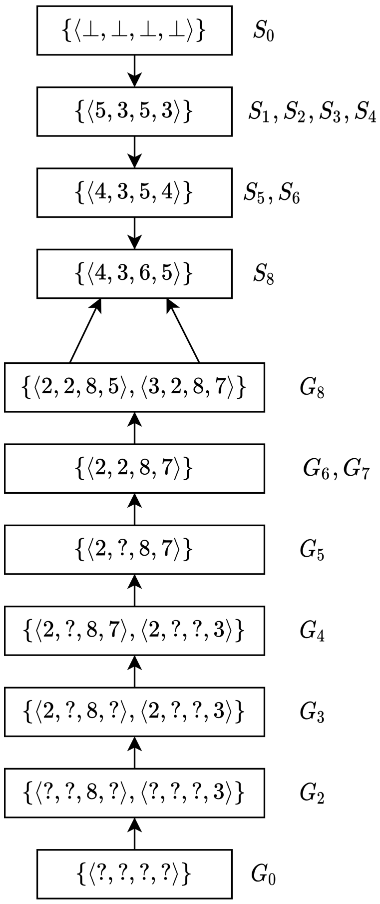
Exercise 3

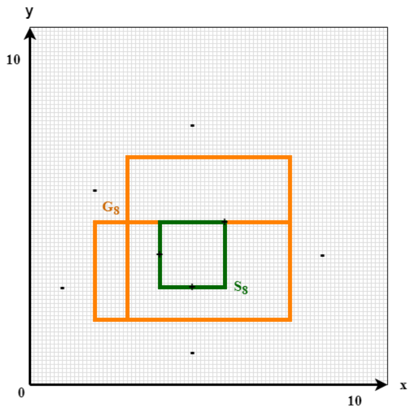
(a) The most general hypothesis $g_0 = \langle 1, 1, 10, 10 \rangle$ corresponds to a square that covers all possible examples.

(b) ☐ $\langle 1, 2, 3, 4 \rangle \geq_g \langle 1, 1, 4, 4 \rangle$
☒ $\langle 2, 3, 6, 7 \rangle \geq_g \langle 3, 4, 5, 7 \rangle$
☐ $\langle 1, 1, 2, 8 \rangle \geq_g \langle 1, 1, 3, 3 \rangle$
☐ $\langle 3, 3, 9, 9 \rangle \geq_g \langle 1, 1, 1, 1 \rangle$

(c) $h_1 = \langle 3, 3, 5, 7 \rangle$
$h_2 = \langle 2, 3, 5, 6 \rangle$

(d) The results of the candidate elimination algorithm are shown in the following:





(e) In this case both S and G would be empty. This is because after Candidate Elimination Algorithm all g in G are first deleted, which are not consistent with x_9. In this case, this affects both g in G. Since s in S is also inconsistent with x_9, it is also deleted. Theoretically, a minimally generalized new s would now have to be found, which is consistent with x_9 but more specific than any g in G. As G is empty, there is no hypothesis for s, which fulfills these requirements and therefore S is empty too.

(f) The rote learner algorithm has no inductive bias because it has no generalization[1]. It stores every examples and classifies x only if it matches with previously observed examples. This means for every positive example the rote learner adds a hypothesis which is only true for the example. If for example point $p(5,3)$ has class 1 then the rote learner will construct the hypothesis $\langle 5,3,5,3\rangle$ and add it to the set of hypothesis.

Exercise 4

(a) Learn(D,1), $\pi = 1$

$$
\begin{array}{lll}
(\mathbf{x}_0, c(\mathbf{x}_0)) = ((1,1),1) & \Rightarrow \mathbf{x}_{0_1} = c(\mathbf{x_0}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1 \\
(\mathbf{x}_1, c(\mathbf{x}_1)) = ((-1,1),-1) & \Rightarrow \mathbf{x}_{1_1} = c(\mathbf{x_1}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1 \\
(\mathbf{x}_2, c(\mathbf{x}_2)) = ((1,1),-1) & \Rightarrow \mathbf{x}_{2_1} \neq c(\mathbf{x_2}) & \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1 \\
(\mathbf{x}_3, c(\mathbf{x}_3)) = ((1,1),1) & \Rightarrow \mathbf{x}_{3_1} = c(\mathbf{x_3}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1 \\
(\mathbf{x}_4, c(\mathbf{x}_4)) = ((1,1),-1) & \Rightarrow \mathbf{x}_{4_1} \neq c(\mathbf{x_4}) & \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1 \\
(\mathbf{x}_5, c(\mathbf{x}_5)) = ((-1,1),1) & \Rightarrow \mathbf{x}_{5_1} \neq c(\mathbf{x_5}) & \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1 \\
(\mathbf{x}_6, c(\mathbf{x}_6)) = ((1,1),-1) & \Rightarrow \mathbf{x}_{6_1} \neq c(\mathbf{x_6}) & \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1 \\
(\mathbf{x}_7, c(\mathbf{x}_7)) = ((-1,-1),-1) & \Rightarrow \mathbf{x}_{7_1} = c(\mathbf{x_7}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1 \\
(\mathbf{x}_8, c(\mathbf{x}_8)) = ((1,1),1) & \Rightarrow \mathbf{x}_{8_1} = c(\mathbf{x_8}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1 \\
(\mathbf{x}_9, c(\mathbf{x}_9)) = ((1,-1),1) & \Rightarrow \mathbf{x}_{9_1} = c(\mathbf{x_9}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1
\end{array}
$$

$\mathcal{L}_- = 6, \quad \mathcal{L}_+ = 4 \Rightarrow w = 1 \Rightarrow y_1(\mathbf{x}) = 1 \cdot \mathbf{x}_1$

$\mathrm{Err}_{tr}(y) = \frac{|\{(\mathbf{x}, c(\mathbf{x})) \in D_{tr} : y(\mathbf{x}) \neq c(\mathbf{x}))\}|}{|D_{tr}|}$

$$
\begin{array}{lll}
y_1(\mathbf{x}_0) = 1 = c(\mathbf{x}_0) & y_1(\mathbf{x}_1) = -1 = c(\mathbf{x}_0) & y_1(\mathbf{x}_2) = 1 \neq c(\mathbf{x}_2) \\
y_1(\mathbf{x}_3) = 1 = c(\mathbf{x}_3) & y_1(\mathbf{x}_4) = 1 \neq c(\mathbf{x}_4) & y_1(\mathbf{x}_5) = -1 \neq c(\mathbf{x}_5) \\
y_1(\mathbf{x}_6) = 1 \neq c(\mathbf{x}_6) & y_1(\mathbf{x}_7) = -1 = c(\mathbf{x}_7) & y_1(\mathbf{x}_8) = 1 = c(\mathbf{x}_8) \\
y_1(\mathbf{x}_9) = 1 = c(\mathbf{x}_9) & &
\end{array}
$$

[1] http://ml.informatik.uni-freiburg.de/former/_media/documents/teaching/ss12/ml/02_conceptlearning.printer.pdf

$\Rightarrow \text{Err}_{tr}(y) = \frac{4}{10}$

(b) Learn($D \setminus D_{test}$,1), $\pi = 1$

$$\begin{aligned}
&(\mathbf{x}_0, c(\mathbf{x}_0)) = ((1,1),1) && \Rightarrow \mathbf{x}_{0_1} = c(\mathbf{x_0}) && \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1\\
&(\mathbf{x}_1, c(\mathbf{x}_1)) = ((-1,1),-1) && \Rightarrow \mathbf{x}_{1_1} = c(\mathbf{x_1}) && \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1\\
&(\mathbf{x}_2, c(\mathbf{x}_2)) = ((1,1),-1) && \Rightarrow \mathbf{x}_{2_1} \neq c(\mathbf{x_2}) && \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1\\
&(\mathbf{x}_3, c(\mathbf{x}_3)) = ((1,1),1) && \Rightarrow \mathbf{x}_{3_1} = c(\mathbf{x_3}) && \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1\\
&(\mathbf{x}_4, c(\mathbf{x}_4)) = ((1,1),-1) && \Rightarrow \mathbf{x}_{4_1} \neq c(\mathbf{x_4}) && \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1\\
&(\mathbf{x}_5, c(\mathbf{x}_5)) = ((-1,1),1) && \Rightarrow \mathbf{x}_{5_1} \neq c(\mathbf{x_5}) && \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1\\
&(\mathbf{x}_6, c(\mathbf{x}_6)) = ((1,1),-1) && \Rightarrow \mathbf{x}_{6_1} \neq c(\mathbf{x_6}) && \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1
\end{aligned}$$

$\mathcal{L}_- = 3, \quad \mathcal{L}_+ = 4 \Rightarrow w = -1 \Rightarrow y_1^{'}(\mathbf{x}) = -1 \cdot \mathbf{x}_1$

$\text{Err}(y, D_{test}) = \frac{|\{(\mathbf{x}, c(\mathbf{x})) \in D_{test} : y^{'}(\mathbf{x}) \neq c(\mathbf{x}))\}|}{|D_{test}|}$

$$y_1^{'}(\mathbf{x}_7) = 1 \neq c(\mathbf{x}_7) \quad y_1^{'}(\mathbf{x}_8) = -1 \neq c(\mathbf{x}_8) \qquad y_1^{'}(\mathbf{x}_9) = -1 \neq c(\mathbf{x}_9)$$

$\Rightarrow \text{Err}(y, D_{test}) = \frac{3}{3} = 1$

(c) $k = 2, \quad D_{val_1} = \{0,1,2,3\}, \quad D_{val_2} = \{4,5,6\}, \quad D_{test} = \{7,8,9\}$

$y_{i_{\pi_j}}^{'}$ is the classifier trained on $D_{tr} = D \setminus (D_{test} \cup D_{val_i})$ for the hyperparameter π_j.

$\pi^* = \underset{\pi_j, j=1,\ldots,m}{\text{argmin}} \sum_{i=1}^{k} Err\left(y_{i_{\pi_j}}^{'}, D_{val_i}\right).$

Training on validation set 1.

$D_{tr_1} = D \setminus (D_{test} \cup D_{val_1}) = \{4,5,6\}$
Learn(D_{tr_1}, 1)

$$\begin{aligned}
&(\mathbf{x}_4, c(\mathbf{x}_4)) = ((1,1),-1) && \Rightarrow \mathbf{x}_{4_1} \neq c(\mathbf{x_4}) && \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1\\
&(\mathbf{x}_5, c(\mathbf{x}_5)) = ((-1,1),1) && \Rightarrow \mathbf{x}_{5_1} \neq c(\mathbf{x_5}) && \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1\\
&(\mathbf{x}_6, c(\mathbf{x}_6)) = ((1,1),-1) && \Rightarrow \mathbf{x}_{6_1} \neq c(\mathbf{x_6}) && \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1
\end{aligned}$$

$\mathcal{L}_- = 0, \quad \mathcal{L}_+ = 3 \Rightarrow w = -1 \Rightarrow y_{1_{\pi_1}}^{'}(\mathbf{x}) = -1 \cdot \mathbf{x}_1$

$\text{Err}(y_{1_{\pi_1}}^{'}, D_{val_1}) = \frac{|\{(\mathbf{x}, c(\mathbf{x})) \in D_{val_1} : y_{1_{\pi_1}}^{'}(\mathbf{x}) \neq c(\mathbf{x}))\}|}{|D_{val_1}|}$

$$\begin{aligned}
&y_{1_{\pi_1}}^{'}(\mathbf{x}_0) = -1 \neq c(\mathbf{x}_0) \quad y_{1_{\pi_1}}^{'}(\mathbf{x}_1) = 1 \neq c(\mathbf{x}_1) \qquad y_{1_{\pi_1}}^{'}(\mathbf{x}_2) = -1 = c(\mathbf{x}_2)\\
&y_{1_{\pi_1}}^{'}(\mathbf{x}_3) = -1 \neq c(\mathbf{x}_3)
\end{aligned}$$

$\Rightarrow \text{Err}(y_{1_{\pi_1}}^{'}, D_{val_1}) = \frac{3}{4}$

Learn(D_{tr_1}, 2)

$$\begin{aligned}
&(\mathbf{x}_4, c(\mathbf{x}_4)) = ((1,1),-1) && \Rightarrow \mathbf{x}_{4_2} \neq c(\mathbf{x_4}) && \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1\\
&(\mathbf{x}_5, c(\mathbf{x}_5)) = ((-1,1),1) && \Rightarrow \mathbf{x}_{5_2} = c(\mathbf{x_5}) && \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1\\
&(\mathbf{x}_6, c(\mathbf{x}_6)) = ((1,1),-1) && \Rightarrow \mathbf{x}_{6_2} \neq c(\mathbf{x_6}) && \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1
\end{aligned}$$

$\mathcal{L}_- = 1, \quad \mathcal{L}_+ = 2 \Rightarrow w = -1 \Rightarrow y^{'}_{1_{\pi_2}}(\mathbf{x}) = -1 \cdot \mathbf{x}_2$

$\mathrm{Err}(y^{'}_{1_{\pi_2}}, D_{val_1}) = \frac{|\{(\mathbf{x}, c(\mathbf{x})) \in D_{val_1} : y^{'}_{1_{\pi_2}}(\mathbf{x}) \neq c(\mathbf{x}))\}|}{|D_{val_1}|}$

$$y^{'}_{1_{\pi_2}}(\mathbf{x}_0) = -1 \neq c(\mathbf{x}_0) \quad y^{'}_{1_{\pi_2}}(\mathbf{x}_1) = -1 = c(\mathbf{x}_1) \quad y^{'}_{1_{\pi_2}}(\mathbf{x}_2) = -1 = c(\mathbf{x}_2)$$
$$y^{'}_{1_{\pi_2}}(\mathbf{x}_3) = -1 \neq c(\mathbf{x}_3)$$

$\Rightarrow \mathrm{Err}(y^{'}_{1_{\pi_2}}, D_{val_1}) = \frac{2}{4}$

Training on validation set 2.

$D_{tr_2} = D \setminus (D_{test} \cup D_{val_2}) = \{0, 1, 2, 3\}$
$\mathrm{Learn}(D_{tr_2}, 1)$

$$\begin{array}{lll}
(\mathbf{x}_0, c(\mathbf{x}_0)) = ((1,1),1) & \Rightarrow \mathbf{x}_{0_1} = c(\mathbf{x_0}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1 \\
(\mathbf{x}_1, c(\mathbf{x}_1)) = ((-1,1),-1) & \Rightarrow \mathbf{x}_{1_1} = c(\mathbf{x_1}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1 \\
(\mathbf{x}_2, c(\mathbf{x}_2)) = ((1,1),-1) & \Rightarrow \mathbf{x}_{2_1} \neq c(\mathbf{x_2}) & \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1 \\
(\mathbf{x}_3, c(\mathbf{x}_3)) = ((1,1),1) & \Rightarrow \mathbf{x}_{3_1} = c(\mathbf{x_3}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1
\end{array}$$

$\mathcal{L}_- = 3, \quad \mathcal{L}_+ = 1 \Rightarrow w = 1 \Rightarrow y^{'}_{2_{\pi_1}}(\mathbf{x}) = 1 \cdot \mathbf{x}_1$

$\mathrm{Err}(y^{'}_{2_{\pi_1}}, D_{val_2}) = \frac{|\{(\mathbf{x}, c(\mathbf{x})) \in D_{val_2} : y^{'}_{2_{\pi_1}}(\mathbf{x}) \neq c(\mathbf{x}))\}|}{|D_{val_2}|}$

$$y^{'}_{2_{\pi_1}}(\mathbf{x}_4) = 1 \neq c(\mathbf{x}_0) \quad y^{'}_{2_{\pi_1}}(\mathbf{x}_5) = -1 \neq c(\mathbf{x}_1) \quad y^{'}_{2_{\pi_1}}(\mathbf{x}_6) = 1 \neq c(\mathbf{x}_2)$$

$\Rightarrow \mathrm{Err}(y^{'}_{2_{\pi_1}}, D_{val_2}) = \frac{3}{3}$

$D_{tr_2} = D \setminus (D_{test} \cup D_{val_2}) = \{0, 1, 2, 3\}$
$\mathrm{Learn}(D_{tr_2}, 2)$

$$\begin{array}{lll}
(\mathbf{x}_0, c(\mathbf{x}_0)) = ((1,1),1) & \Rightarrow \mathbf{x}_{0_2} = c(\mathbf{x_0}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1 \\
(\mathbf{x}_1, c(\mathbf{x}_1)) = ((-1,1),-1) & \Rightarrow \mathbf{x}_{1_2} \neq c(\mathbf{x_1}) & \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1 \\
(\mathbf{x}_2, c(\mathbf{x}_2)) = ((1,1),-1) & \Rightarrow \mathbf{x}_{2_2} \neq c(\mathbf{x_2}) & \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1 \\
(\mathbf{x}_3, c(\mathbf{x}_3)) = ((1,1),1) & \Rightarrow \mathbf{x}_{3_2} = c(\mathbf{x_3}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1
\end{array}$$

$\mathcal{L}_- = 2, \quad \mathcal{L}_+ = 2 \Rightarrow w = 1 \Rightarrow y^{'}_{2_{\pi_2}}(\mathbf{x}) = 1 \cdot \mathbf{x}_1$

$\mathrm{Err}(y^{'}_{2_{\pi_2}}, D_{val_2}) = \frac{|\{(\mathbf{x}, c(\mathbf{x})) \in D_{val_2} : y^{'}_{2_{\pi_2}}(\mathbf{x}) \neq c(\mathbf{x}))\}|}{|D_{val_2}|}$

$$y^{'}_{2_{\pi_2}}(\mathbf{x}_4) = 1 \neq c(\mathbf{x}_0) \quad y^{'}_{2_{\pi_2}}(\mathbf{x}_5) = 1 = c(\mathbf{x}_1) \quad y^{'}_{2_{\pi_2}}(\mathbf{x}_6) = 1 \neq c(\mathbf{x}_2)$$

$\Rightarrow \mathrm{Err}(y^{'}_{2_{\pi_2}}, D_{val_2}) = \frac{2}{3}$

Computing π^*

$\mathrm{Err}(y^{'}_{1_{\pi_1}}, D_{val_1}) = \frac{3}{4}$
$\mathrm{Err}(y^{'}_{1_{\pi_2}}, D_{val_1}) = \frac{2}{4}$
$\mathrm{Err}(y^{'}_{2_{\pi_1}}, D_{val_2}) = \frac{3}{3}$
$\mathrm{Err}(y^{'}_{2_{\pi_2}}, D_{val_2}) = \frac{2}{3}$

$\pi_1 : \mathrm{Err}(y^{'}_{1_{\pi_1}}, D_{val_1}) + \mathrm{Err}(y^{'}_{2_{\pi_1}}, D_{val_2}) = \frac{21}{12}$

$\pi_2 : \mathrm{Err}(y^{'}_{1_{\pi_2}}, D_{val_1}) + \mathrm{Err}(y^{'}_{2_{\pi_2}}, D_{val_2}) = \frac{14}{12}$

$\Rightarrow \pi^* = \pi_2$ since the sum of errors on all validation sets if lower for π_2 then for π_1.

Learn($D \setminus D_{test}$,2), $\pi = 2$

$$
\begin{array}{lll}
(\mathbf{x}_0, c(\mathbf{x}_0)) = ((1,1),1) & \Rightarrow \mathbf{x}_{0_1} = c(\mathbf{x_0}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1 \\
(\mathbf{x}_1, c(\mathbf{x}_1)) = ((-1,1),-1) & \Rightarrow \mathbf{x}_{1_1} \neq c(\mathbf{x_1}) & \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1 \\
(\mathbf{x}_2, c(\mathbf{x}_2)) = ((1,1),-1) & \Rightarrow \mathbf{x}_{2_1} \neq c(\mathbf{x_2}) & \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1 \\
(\mathbf{x}_3, c(\mathbf{x}_3)) = ((1,1),1) & \Rightarrow \mathbf{x}_{3_1} = c(\mathbf{x_3}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1 \\
(\mathbf{x}_4, c(\mathbf{x}_4)) = ((1,1),-1) & \Rightarrow \mathbf{x}_{4_1} \neq c(\mathbf{x_4}) & \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1 \\
(\mathbf{x}_5, c(\mathbf{x}_5)) = ((-1,1),1) & \Rightarrow \mathbf{x}_{5_1} = c(\mathbf{x_5}) & \Rightarrow \mathcal{L}_- = \mathcal{L}_- + 1 \\
(\mathbf{x}_6, c(\mathbf{x}_6)) = ((1,1),-1) & \Rightarrow \mathbf{x}_{6_1} \neq c(\mathbf{x_6}) & \Rightarrow \mathcal{L}_+ = \mathcal{L}_+ + 1
\end{array}
$$

$\mathcal{L}_- = 3, \quad \mathcal{L}_+ = 4 \Rightarrow w = -1 \Rightarrow y^{'}_{\pi^*}(\mathbf{x}) = -1 \cdot \mathbf{x}_2$

$\mathrm{Err}(y_{\pi^*}, D_{test}) = \frac{|\{(\mathbf{x},c(\mathbf{x})) \in D_{test} : y^{'}_{\pi^*}(\mathbf{x}) \neq c(\mathbf{x}))\}|}{|D_{test}|}$

$$y^{'}_{\pi^*}(\mathbf{x}_7) = 1 \neq c(\mathbf{x}_7) \qquad y^{'}_{\pi^*}(\mathbf{x}_8) = -1 \neq c(\mathbf{x}_8) \qquad y^{'}_{\pi^*}(\mathbf{x}_9) = 1 = c(\mathbf{x}_9)$$

$\Rightarrow \mathrm{Err}(y_{\pi^*}, D_{test}) = \frac{2}{3}$