

Assignment

Machine Learning

1 Apply Candidate - Elimination Algorithm for the following dataset.

Ex	origin	Manufacturer	Color	Decade	Type	Target Concept
1	Japan	Honda	Blue	1980	Economy	+
2	Japan	Toyota	Green	1970	Sports	-
3	Japan	Toyota	Blue	1990	Economy	+
4	USA	Chrysler	Red	1980	Economy	-
5	Japan	Honda	White	1980	Economy	+

∴ Firstly Initialize G_0 and S_0

$$S_0 = (\emptyset, \emptyset, \emptyset, \emptyset, \emptyset)$$

// First Instance

$$S_1 = (\text{Japan}, \text{Honda}, \text{Blue}, 1980, \text{Economy})$$

// Second Instance

$$S_2 = S_1$$

// Second Instance

$$S_3 = (\text{Japan}, ?, \text{Blue}, ?, \text{Economy})$$

// Third Instance

$$S_4 = (\text{Japan}, ?, \text{Blue}, ?, \text{Economy})$$

// Fourth Instance

$$S_5 = (\text{Japan}, ?, ?, ?, \text{Economy})$$

// Fifth Instance

$$G_5 = (\text{Japan}, ?, ?, ?, \text{Economy})$$

$$G_4 = (?, ?, \text{Blue}, ?, ?) (\text{Japan}, ?, ?, ?, \text{Economy})$$

$$G_4 = (?, ?, \text{Blue}, ?, ?) (\text{Japan}, ?, ?, ?, \text{Economy}) (\text{USA}, ?, ?, ?, \text{Economy})$$

$$G_4 = (?, ?, \text{Blue}, ?, ?) (\text{Japan}, ?, ?, ?, \text{Economy}) (\text{USA}, ?, ?, ?, \text{Economy})$$

$$G_4 = (?, \text{Honda}, ?, ?, \text{Economy}) (?, \text{Chrysler}, ?, ?, \text{Economy}) (?, \text{Toyota}, ?, ?, \text{Economy})$$

$$G_4 = (?, ?, \text{Red}, ?, \text{Economy}) (?, ?, \text{White}, ?, \text{Economy}) (?, ?, ?, 1980, \text{Economy})$$

$$G_4 = (?, ?, ?, 1970, \text{Economy}) (?, ?, ?, 1990, \text{Economy}) (?, ?, ?, ?, \text{Sports})$$

$$G_3 = (?, ?, \text{Blue}, ?, ?) (?, ?, ?, ?, \text{Economy})$$

$$G_3 = (?, ?, \text{Blue}, ?, ?) (?, ?, \text{Blue}, ?, ?) (?, ?, ?, 1980, ?)$$

$$G_2 = (?, \text{Honda}, ?, ?, ?) (?, ?, \text{Blue}, ?, ?) (?, ?, ?, ?, \text{Economy})$$

$$G_0 = (\text{USA}, ?, ?, ?, ?) (?, \text{Honda}, ?, ?, ?) (?, \text{Chrysler}, ?, ?, ?)$$

$$G_0 = (?, ?, \text{Blue}, ?, ?) (?, ?, \text{Red}, ?, ?) (?, ?, \text{White}, ?, ?)$$

$$G_0 = (?, ?, ?, 1980, ?) (?, ?, ?, 1990, ?) (?, ?, ?, ?, \text{Economy})$$

$$G_1 = G_0$$

$$G_0 = (?, ?, ?, ?, ?)$$

$$G_5 = (\text{Japan}, ?, ?, ?, \text{Economy})$$

$$G_5 = (\text{Japan}, ?, ?, ?, \text{Economy})$$

These are the final set of hypotheses which
Consistent with the training instance.

② Apply List-Then-Eliminate Algorithm to the following
data set to generate the version space.

Ex	Sky	Temperature	Target
1	Sunny	High	Yes
2	Sunny	Normal	Yes
3	Rainy	High	No
4	Rainy	Normal	No

Number of Instance

$$\text{Sky} = 2$$

$$\text{Temperature} = 2$$

$$= 2 \times 2 = 4$$

Syntactically distinct hypothesis

$$(2+2) * (2+2) = 16$$

Semantically distinct hypothesis

$$1 + (3+3) = 10$$

Step 1 Version Space

7

$$h_1 = \langle \text{Sunny}, \text{High} \rangle$$

$$h_2 = \langle \text{Sunny}, \text{Normal} \rangle$$

$$h_3 = \langle \text{Rainy}, \text{High} \rangle$$

$$h_4 = \langle \text{Rainy}, \text{Normal} \rangle$$

$$h_5 = \langle \text{Sunny}, ? \rangle$$

$$h_6 = \langle \text{Rainy}, ? \rangle$$

$$h_7 = \langle ?, \text{High} \rangle$$

$$h_8 = \langle ?, \text{Normal} \rangle$$

$$h_9 = \langle ?, ? \rangle$$

$$h_{10} = \langle \phi, \phi \rangle$$

~~Step 2~~ Constant hypothesis

$$h_1 = \langle \text{Sunny}, \text{High} \rangle$$

$$h_5 = \langle \text{Sunny}, ? \rangle$$

$$h_7 = \langle ?, \text{High} \rangle$$

$$h_9 = \langle ?, ? \rangle$$

$$h_5 = \langle \text{Sunny}, ? \rangle$$

$$h_9 = \langle ?, ? \rangle$$

8. output

$$\langle \text{Sunny}, ? \rangle$$