CIS 678 HW 2

Thomas Bailey

February 2019

1. Find-S Algorithm

$$h_0 = < -, -, -, -, -, - >$$
$$h_1 = < Strong, Warm, Warm, AM, Sunny, Weekend >$$
$$h_2 = < Strong, Warm, Warm, *, Sunny, Weekend >$$
$$h_3 = < Strong, Warm, *, *, Sunny, * >$$

   Hypothesis h_3 is consistent with all negative training instances.

2. Candidate Elimination Algorithm

   (a) i. $G_0 = \{< *, *, *, *, *, * >\}$
   $S_0 = \{< -, -, -, -, -, - >\}$

   ii. The first example is positive, so set S_0 is too specific.
   $G_1 = \{< *, *, *, *, *, * >\}$
   $S_1 = \{< Strong, Warm, Warm, AM, Sunny, Weekend >\}$

   iii. The second example is positive, and set S_1 is not consistent with the new example.
   $G_2 = \{< *, *, *, *, *, * >\}$
   $S_2 = \{< Strong, Warm, Warm, *, Sunny, Weekend >\}$

   iv. The third example is positive, and set S_2 is not consistent with the new example.
   $G_3 = \{< *, *, *, *, *, * >\}$
   $S_3 = \{< Strong, Warm, *, *, Sunny, * >\}$

   v. The fourth example is negative, so G_3 is too general.
   $G_4 = \{< *, Warm, *, *, *, * >, < *, *, *, *, Sunny, * >\}$
   $S_4 = \{< Strong, Warm, *, *, Sunny, * >\}$

   (b) Swapping the data has no effect on the final hypothesis sets.

   (c) The set S_4 contains the hypothesis obtained from the Find-S algorithm.

3. (a) We can classify data A as a good day to fish, because the data is consistent with the results of the Candidate Elimination algorithm. Because it is warm and sunny with strong wind, it is a good day to fish.

   (b) We can classify data B as a bad day to fish. We know it must be warm and sunny with strong wind to fish, and data B isn't consistent with this hypothesis.