

## Problem Set 1

Bangqi Wang

Handed In: February 2, 2017

## 1. Answer to problem 1

- a. The main idea of the algorithm is to construct an initial hypothesis and then update the hypothesis when iterating the training set. For instance, the algorithm constructs an initial complete hypothesis that is the conjunction of all variables and their negations. i.e.  $\mathbf{H} = (x_1 \wedge x_2 \wedge \dots \wedge x_n \wedge \neg x_1 \wedge \neg x_2 \wedge \dots \wedge \neg x_n)$ . Then, for each positive sequences in training set, the algorithm updates the hypothesis by removing the contradictions. i.e.  $x_i = a$  in hypothesis, but  $x_i = b$  in sequence. After finishing iterating positive sequences, using negative sequences to check the hypothesis. If there is any contradiction between negative sequence and hypothesis, the algorithm will indicate that there is no consistent hypothesis.

---

**Algorithm 1** Pseudocode:

---

```

1: procedure MYPROCEDURE
2:    $result \leftarrow (x_1 = 1) \wedge (x_2 = 1) \wedge \dots \wedge (x_n = 1) \wedge (x_1 = 0) \wedge (x_2 = 0) \wedge \dots \wedge (x_n = 0)$ 
3:   for each '+' sequence  $S$  in training set do
4:     if  $(x_i = v) \neq S_i$  then
5:       remove  $(x_i = v)$  from  $result$ 
6:     end if
7:   end for
8:   for each '-' sequence  $S$  in training set do
9:     if  $S$  contradict with  $result$  then
10:      Indicate: no consistent hypothesis
11:      halt
12:    end if
13:   end for
14:   return  $result$ 
15: end procedure

```

---

- b. Inorder to prove the correctness of this algorithm, I will prove that there are neither **false positive** nor **false negative**.
- b.1 If there is a **false positive**, there are at least one condition absent from  $result$ , say  $(x_i = v)$ . However, in the first loop, the algorithm eliminated  $(x_i = v)$  from complete conjunction, because it contradicts with other positive examples. Those two assumptions are mutually conflicting.
- b.2 If there is a **false negative**, there are at least one incorrect condition in  $result$ , say  $(x_i = v)$ . However, in the first loop, the algorithm eliminated all conditions that contradict with other positive examples. The remaining

$(x_i = v)$  should not contradict with any positive sequence. Therefore, those two assumptions are mutually conflicting.

Therefore, the *result* contains neither **false positive** nor **false negative**. The algorithm is correct.

- c. The algorithm traverses  $m$  training examples and checks each example with hypothesis *result* that has  $2n$  conditions at worst. The runtime should be  $\mathbf{O}(mn)$ .
- d. There might be **false negative** in this situation. The example labeled as positive must be positive. However, the example labeled as negative might be positive, because the algorithm updates the hypothesis according to positive examples in training examples. The hypothesis changed from general to specific. The hypothesis *result* is the superset of real hypothesis  $\mathbf{H}$ .

## 2. Answer to problem 2

- a. Assume a vector  $\vec{w}$  that is perpendicular to the hyperplane. Point  $x^* = (x_1^*, x_2^*, \dots, x_n^*)$  stands for any point in hyperplane. Define  $\vec{v}$  as vector from point  $x^0 = (x_0^0, x_1^0, \dots, x_n^0)$  to point  $x^*$ .

$$\vec{v} = [x_0^* - x_0^0, \quad x_1^* - x_1^0, \quad \dots, \quad x_n^* - x_n^0]^T$$

Then, the length  $D$  of the projection of  $\vec{v}$  onto  $\vec{w}$  is the distance from the point to the plane.

$$D = |\text{proj}_{\vec{w}} \vec{v}| \tag{1}$$

$$= \frac{|\vec{w}^T \cdot \vec{v}|}{\|\vec{w}\|} \tag{2}$$

$$= \frac{|w_0 \cdot (x_0^* - x_0^0) + w_1 \cdot (x_1^* - x_1^0) + \dots + w_n \cdot (x_n^* - x_n^0)|}{\sqrt{w_0^2 + w_1^2 + \dots + w_n^2}} \tag{3}$$

$$= \frac{|w_0 \cdot x_0^* - w_0 \cdot x_0^0 + w_1 \cdot x_1^* - w_1 \cdot x_1^0 + \dots + w_n \cdot x_n^* - w_n \cdot x_n^0|}{\|\vec{w}\|} \tag{4}$$

$$= \frac{|\theta + w_0 \cdot x_0^* + w_1 \cdot x_1^* + \dots + w_n \cdot x_n^*|}{\|\vec{w}\|} \tag{5}$$

$$= \frac{|\vec{w}^T \cdot \vec{x}^* + \theta|}{\|\vec{w}\|} \tag{6}$$

- b.

$$\vec{w}^T \cdot \vec{x} + \theta_2 = 0 \tag{7}$$

$$\vec{w}^T \cdot \vec{x} = -\theta_2 \tag{8}$$

$$D = \frac{|\vec{w}^T \cdot \vec{x}^* + \theta|}{\|\vec{w}\|} \tag{9}$$

$$= \frac{|\theta_1 - \theta_2|}{\|\vec{w}\|} \tag{10}$$

## 3. Answer to problem 3

- a. a.1
- a.2
- a.3
- b. b.1
- b.2
- b.3
- b.4