ĐỀ 1

Câu 1:

a.

|  |  |  |  |
| --- | --- | --- | --- |
| **Nút** | **h** | **i** | **H ( lấy giá trị nhỏ nhất )** |
| S | 6 | 9 | 7 |
| A | 3 | 7 | 9 |
| B | 6 | 5 | 6 |
| C | 5 | 6 | 6 |
| D | 4 | 7 | 6 |
| E | 8 | 7 | 8 |
| F | 3 | 3 | 3 |
| H | 1 | 2 | 1 |
| I | 2 | 2 | 2 |
| G | 0 | 0 | 0 |

h là hàm chấp nhận được.

b.

 A\* theo h

|  |  |  |
| --- | --- | --- |
| Bước | Nút mở rộng | Nút, hàm f |
| 1 |  | S (f=6) |
| 2 | S | A\_S (f=4), B\_S (f=7) |
| 3 | A\_S | B\_S(f=7), C\_A(f=10), D\_A(f=8) |
| 4 | B\_S | C\_A(f=10), D\_A(f=8), E\_B(11), F\_B(7) |
| 5 | F\_B | C\_A(f=10), D\_A(f=8), E\_B(11), I\_F(7) |
| 6 | I\_F | C\_A(f=10), D\_A(f=8), E\_B(11), H\_I (f=7) |
| 7 | H\_I | C\_A(f=10), D\_A(f=8), E\_B(11), G\_H(f=7) |
| 8 | G\_H | Dừng |

Chi phí : G <- H <-I <- F <- B <- S ( chí phí 7)

A\* theo i

|  |  |  |
| --- | --- | --- |
| Bước | Nút mở rộng | Nút, hàm f |
| 1 |  | S (f=9) |
| 2 | S | A\_S (f=8), B\_S (f=6) |
| 3 | B\_S | A\_S (f=8), E\_B(10), F\_B(7) |
| 4 | F\_B | A\_S (f=8), E\_B(10), I\_F (7) |
| 5 | I\_F | A\_S (f=8), E\_B(10), H\_I (8) |
| 6 | A\_S | C\_A(f=11), D\_A(f=11),  E\_B(10), H\_I (f=8) |
| 7 | H\_I | C\_A(f=11), D\_A(f=11), E\_B(10),  G\_H(f=7) |
| 8 | G\_H | Dừng |

Câu 2:

Câu 3:

b.  ( đề 1) P(V) = 0.2

So sánh P(V|C) và P(-V|C)

P(V|C) = P(C|V) . P(V) / P(C)  ( Định lý bayes)

Từ mạng bayes và các xs đk đã cho: P(V|C) = 0.7 . 0.2 / P (C)

Tương tự : P (-V|C ) = 0.16 / P(C)

P(V|C) + P(-V|C) = 1

= > P (-V|C) > P(V|C) => máy k nhiễm virut

Câu 4.

a. Tìm nút gốc cho cây quyết định :

- Tính information gain  của từng thuộc tính với tập dữ liệu đã có .

- Chọn thuộc tính có information gain lớn nhất

Gain (S, A) = 0, Gain (S, B) = 0.251, Gain(S, C) = 0.459

=> C làm nút gốc

b.

+  Sử dụng hàm kc euclid

+ Tính kc của z tới từng ví dụ  trong bảng đã cho

+ Xác định k láng giềng gần nhất,  nhãn E hồi quy là TB cộng các nhãn của K láng giềng đó

z = (A=2, B=2, C=1)  => Nhãn E = 1.75