Midterm Decision Tree

Monday, March 15, 2021 6:56 PM

Attributes = 4, 5, 6, 7, 8, 9 Target = 10

(1)

Column 4

(1) $H(1) = 0$

(2) $H(2) = -(1/2)\log_2(1/2) - (1/2)\log_2(1/2) = 1$

(2) $H(3) = 1$

(4) $H(4) = 1$

(1) $H(5) = 0$

$E(C4) = 0(1/10) + 1(2/10) + 1(2/10) + 1(4/10) + 0(1/10) = 0.8$

Column 5

(2) $H(1) = 0$

(0) $H(2) = 0$

(3) $H(3) = (1/3)\log_2(1/3) - (2/3)\log_2(2/3) = 0.91829$

(3) $H(4) = H(3) = 0.91829$

(2) $H(5) = 0$

$E(C5) = 0(2/10) + 0.91829(3/10) + 0.91829(3/10) + 0(2/10) = 0.5509$

Column 6

(5) $H(2) = -(4/5)\log_2(4/5) - (1/5)\log_2(1/5) = 0.7219$

(1) $H(3) = 0$

(4) $H(5) = -(3/4)\log_2(3/4) - (1/4)\log_2(1/4) = 0.811278$

$E(C6) = 0.7219(5/10) + 0.811278(4/10) = 0.68546$

Column 7

(4) $H(\emptyset) = 1$

(6) $H(1) = 1$

$E(C7) = 1(0.4) + 1(0.6) = 1$

Column 8

(3) $H(\emptyset) = 0.91829$

(7) $H(1) = -(3/7)\log_2(3/7) - (4/7)\log_2(4/7) = 0.985228$

$E(C8) = 0.91829(3/10) + 0.985228(7/10) = 0.96514$

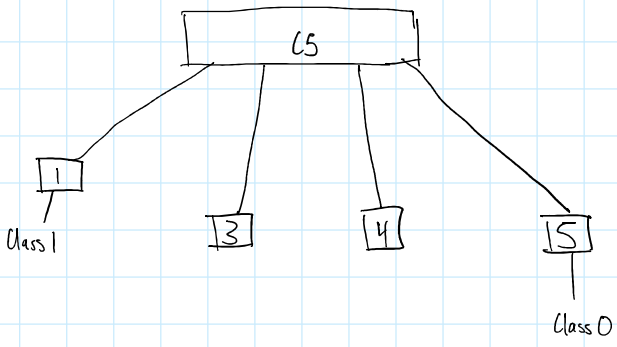
Column 9

(7) $H(\emptyset) = -(2/7)\log_2(2/7) - (5/7)\log_2(5/7) = 0.863120$

(3) $H(1) = 0$

$E(C7) = 0.863120(7/10) = 0.6041843$

∴ based on entropy select column 5 first



E[C4 | C5 = 3] (3)

(1) $H(1) = 0$

(1) $H(4) = 0$

(1) $H(5) = 0$

E[C6 | C5 = 3] (3)

(2) $H(2) = 0$

(1) $H(5) = 0$

E[C7 | C5 = 3] (3)

(1) $H(1) = 0$

(2) $H(0) = 0$

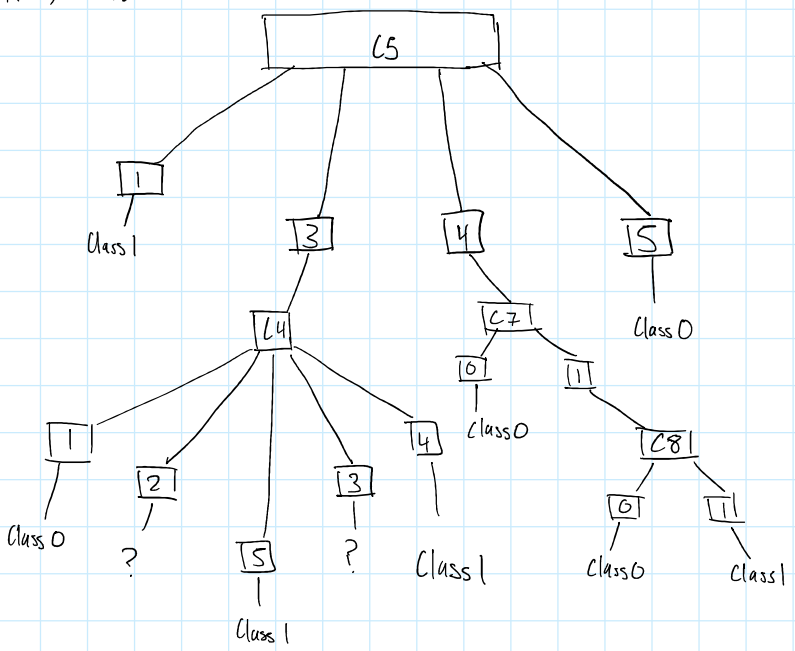
$E(C4) = 0$ $E(C6) = 0$ $E(C7) = 0$

$E[C8|C5=3]$ $E[C9|C5=3] = 0$

(1) (2)

$H(0) = 0$ $H(1) = 1$ *Select C4

$E(C8) = 1(0.66) = 0.\bar{6}$



$E[C6|C5=4]$ (3) $E[C7|C5=4] = E(C6) = 0.\bar{6}$

(1) (2)

$H(3) = 0$ $H(2) = 1$ $E[C8|C5=4] = 0.\bar{6}$

$E(C6) = 0.\bar{6}$ $E[C9|C5=4]$

$H(0) = -(2/3)\log_2(2/3) - (1/3)\log_2(1/3) = 0.91829$

*Slect C7

$E[C6|C7=1] = E[C9|C7=1]$

$E[C8|C7=1] = 0$

Errors = 0/10

Q3 - Bayesian

Column 4

Class ∅

$P(x=4 \mid C10=\emptyset) = 2/5 = 0.4$ $P(x=2 \mid C10=\emptyset) = 1/5 = 0.2$ $P(x=1 \mid C10=\emptyset) = 1/5 = 0.2$

$P(x=3 \mid C10=\emptyset) = 1/5 = 0.2$

Class 1

$P(x=2 \mid C10=1) = 1/5 = 0.2$ $P(x=3 \mid C10=1) = 0.2$ $P(x=4 \mid C10=1) = 0.4$

$P(x=5 \mid C10=1) = 0.2$

Column 5

Class ∅

$P(x=4) = 0.4$ $P(x=5) = 0.4$ $P(x=3) = 0.2$

Class 1

$P(x=1) = 0.4$ $P(x=4) = 0.2$ $P(x=3) = 0.4$

Column 6

Class ∅

$P(x=3) = 0.2$ $P(x=5) = 0.6$ $P(x=2) = 0.2$

Class 1

$P(x=2) = 0.8$ $P(x=5) = 0.2$

Column 7

Class ∅

$P(x=0) = 0.4$ $P(x=1) = 0.6$

Class 1

$P(x=0) = 0.4$ $P(x=1) = 0.6$

Column 8

Class 0

$P(x=0) = 0.4$ $P(x=1) = 0.6$

Class1

$P(x=0) = 0.2$ $P(x=1) = 0.8$

Column 9

Class 0

$P(x=0) = 0.4$ $P(x=1) = 0.6$

Class1

$P(x=0) = 1$

Column 1

$\mu_0 = 0.686323$ $\sigma_0 = 0.246122$

$\mu_1 = 0.435551$ $\sigma_1 = 0.274458$

Column 2

$\mu_0 = 0.505554$ $\sigma_0 = 0.2597931$

$\mu_1 = 0.510482$ $\sigma_1 = 0.353941$

Column 3

$\mu_0 = 0.580936$ $\mu_1 = 0.27101$

$\sigma_0 = 0.394454$ $\sigma_1 = 0.128953$

Combined Probability (Numeric)

1) Class 0 = 0.0015362 Class 1: 0.0007749

2) Class 0 = 0.001687 Class 1: 0.000798

3) Class 0 = 0.0016427 Class 1 : 0.000794

4) Class 0 = 0.001588 Class 1: 0.0007646

5) Class 0 = 0.001571137 Class 1: 0.000797

6) Class 0 = 0.001658 Class 1: 0.

1) [4, 4, 3, 0, 1, 0]

$P(X=0) = 0.0015362(0.4)(0.4)(0.2)(0.4)(0.6)(0.4) = 0.0471 \times 10^{-3}$ ∴ Predict class 0

$P(X=1) = 0.0007749(0.6)(0.4)(0.6)(0.4)(0.4) = 0.000178$

2) [25, 5, 0, 1, 1]

$P(X=0) = 0.001687(0.4)(0.6)(0.6)(0.2)(0.4)(0.6) = 11.66 \times 10^{-6}$ Predict class 0

$P(X=1) = 0.000798(0.4)(0.8)(0.2)(0.2) = 10.21 \times 10^{-6}$

3) [4, 5, 5, 1, 0, 1]

$P(X=0) = 0.0016427(0.4)(0.4)(0.6)(0.6)(0.4)(0.6) = 22.71 \times 10^{-6}$

$P(X=1) = 0.000798(0.4)(0.2)(0.6)(0.2)(1) = 7.58 \times 10^{-6}$

4) [1, 3, 5, 1, 1, 1]

$P(X=0) = 0.001588(0.2)(0.2)(0.6)(0.6)(0.6)(0.6) = 8.23 \times 10^{-6}$ Predict class 1

$P(X=1) = 0.0007646(0.4)(0.2)(0.6)(0.8) = 29.36 \times 10^{-6}$

5) [3, 4, 2, 1, 0, 0]

$P(x=0) = 0.0600157113737(0.2)(0.4)(0.2)(0.6)(0.4)(0.4) = 0.482 \times 10^{-6}$