AMELIA. CANTIKA. PUTRI. TANJARO
3022110701

PENAMBANGAN DATA

5 Juni 2023

1) Hitunglah secara manual information gain dari data di samping

- information gain (class : online course)
- information gain (class : Education background)
- information gain (class : Work Status)

- online course

↳ jumlah lulus : 8 → total 15 (Y)

jumlah Gagal : 7

$H(Y) = -(^8/_{15} \log_2 {}^8/_{15}) - (^7/_{15} \log_2 {}^7/_{15})$

$= -(0,6 \log_2(0,6)) - (0,3 \log_2(0,3))$

$= 0,13 + 0,15$

$= 0,28$

$= 0,3$ //

$H(N) = -(^3/_7 \log_2 {}^3/_7) - (^4/_7 \log {}^4/_7)$

$= -(0,4 \log_2(0,4)) - (0,6 \log_2(0,6))$

$= 0,16 + 0,13$

$= 0,29$

$= 0,3$

$H(S) = {}^8/_{15}(0,3) + {}^7/_{15}(0,3)$

$= 0,16 + 0,14$

$= 0,3$

- Education Background

jumlah lulus : 8 → total 15 $H(S) = 0,3$

jumlah gagal : 7

- Math

lulus : 4
gagal : 3 > total 7

$H(Ma) = -(4/7 \log_2 4/7) - (3/7 \log 3/7)$

$= -(0,6 \log_2(0,6)) - (0,4 \log_2(0,4))$

$= 0,13 + 0,16$

$= 0,3$

- CS

lulus : 4

gagal : 0

$H(CS) = -(4/4 \log 4/4)$

$= -(1 \log 1)$

$= 0$

- Other

lulus : 6

$H(S(EB) = 7/15(0,3) + 4/15(0) + 4/15(0)$

$= 0,14$

- Work Status

$H(S) = 0,3$

- HW

lulus : 5 gagal : 1

$H(Hw) = (5/6 \log_2 5/6) - (1/6 \log_2 1/6)$

$= 0,07 + 0,13$

$= 0,2$

- w

lulus : 3 , gagal : 6

$H(w) = -(3/9 \log_2 3/9) - (6/9 \log_2 6/9)$

$= 0,15 + 0,13$

$= 0,28$

$= 0,3$

$H(ws) = 6/15 (0,2) + 9/15 (0,3)$

$= 0,08 + 0,18$

$= 0,26$

$= 0,3$

2) Diketahui sebuah problem klasifikasi 3 kelas memiliki confusion matrix sebagai berikut :

- Hitunglah akurasi keseluruhan classifier tersebut
- Berapa presisi class C_2?

Actual class	Predicted Class		
	C_1	C_2	C_3
C_1	110	8	7
C_2	16	130	10
C_3	26	5	120

↳ Benar : 110 + 130 + 120 = 360

Total = 110 + 8 + 7 + 16 + 130 + 10 + 26 + 5 + 120

= 432

Akurasi $360/432 = 0,8$ //

- presisi C_2

prediksi benar $C_2 = 130$

Total prediksi $C_2 = 8 + 130 + 5 = 143$

presisi $\Rightarrow \frac{130}{143} = 0,9$ //