Decision Tree Learning

Studi Kasus

Masalah: Memutuskan apakah akan menunggu meja kosong dalam restoran atau tidak

Atribut:

1. Alternate: is there an alternative restaurant nearby?
2. Bar: is there a comfortable bar area to wait in?
3. Fri/Sat: is today Friday or Saturday?
4. Hungry: are we hungry?
5. Patrons: number of people in the restaurant (None, Some, Full)
6. Price: price range ($, $$, $$$)
7. Raining: is it raining outside?
8. Reservation: have we made a reservation?
9. Type: kind of restaurant (French, Italian, Thai, Burger)
10. WaitEstimate: estimated waiting time (0-10, 10-30, 30-60, >60)

Sampel Data yang akan digunakan sebagai Training Set adalah



Pohon keputusan dibuat berdasarkan ALGORITMA ID3

* ***Rekursi level 0 iterasi 1***

Atribut target = True, False

Berdasarkan atribut target, maka semua kumpulan sampel data = [6+, 6-]

$Entropy$ (S) = $\sum\_{i}^{c}-Pi log\_{2}Pi= -\left(\frac{6}{12} log\_{2}\left(\frac{6}{12}\right)\right)-\left(\frac{6}{12} log\_{2}\left(\frac{6}{12}\right)\right)$ = 1

Kumpulan Atribut = {Alternate, Bar, Fri/Sat, Hungry, Patrons, Price, Rain, Reservation, Type, Wait Estimate}

Kemudian kita akan mencari nilai Information Gain pada tiap atribut yang ada

**Values(Alternate)= True, False**

S = [6+, 6-], |S| = 12

STrue = [3+, 3-], |STrue| = 6

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{3}{6} log\_{2}\left(\frac{3}{6}\right)\right)-\left(\frac{3}{6} log\_{2}\left(\frac{3}{6}\right)\right)$ = 1

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Alternate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{6}{12} Entropy \left(Strue\right)= \frac{6}{12}\*1= \frac{1}{2}$$

SFalse = [3+, 3-], |SFalse| = 6

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{3}{6} log\_{2}\left(\frac{3}{6}\right)\right)-\left(\frac{3}{6} log\_{2}\left(\frac{3}{6}\right)\right)$ = 1

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Alternate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{6}{12} Entropy \left(Sfalse\right)= \frac{6}{12}\*1= \frac{1}{2}$$

Sehingga Information Gain dari Atribut Alternate adalah

Gain(S, Alternate) = Entropy (S) – Reminder (Alternate)

 = $1- \left(\frac{1}{2}+ \frac{1}{2}\right)$

 = 0

**Values(Bar) = True, False**

S = [6+, 6-], |S| = 12

STrue = [3+, 3-], |STrue| = 6

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi= -\left(\frac{3}{6} log\_{2}\left(\frac{3}{6}\right)\right)-\left(\frac{3}{6} log\_{2}\left(\frac{3}{6}\right)\right)$ = 1

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Bar\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{6}{12} Entropy \left(Strue\right)= \frac{6}{12}\*1= \frac{1}{2}$$

SFalse = [3+, 3-], |SFalse| = 6

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi= -\left(\frac{3}{6} log\_{2}\left(\frac{3}{6}\right)\right)-\left(\frac{3}{6} log\_{2}\left(\frac{3}{6}\right)\right)$ = 1

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Bar\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{6}{12} Entropy \left(Sfalse\right)= \frac{6}{12}\*1= \frac{1}{2}$$

Sehingga Information Gain dari Atribut Bar adalah

Gain(S, Bar) = Entropy (S) – Reminder (Bar)

 = $1- \left(\frac{1}{2}+ \frac{1}{2}\right)$

 = 0

**Values(Fri/Sat) = True, False**

S = [6+, 6-], |S| = 12

STrue = [2+, 3-], |STrue| = 5

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{5} log\_{2}\left(\frac{2}{5}\right)\right)-\left(\frac{3}{5} log\_{2}\left(\frac{3}{5}\right)\right)$ = 0,970951

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(\frac{Fri}{Sat}\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{5}{12} Entropy \left(Strue\right)= \frac{5}{12}\*0,970951=0,404563$$

SFalse = [4+, 3-], |SFalse| = 7

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{4}{7} log\_{2}\left(\frac{4}{7}\right)\right)-\left(\frac{3}{7} log\_{2}\left(\frac{3}{7}\right)\right)$ = 0,985228

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(\frac{Fri}{Sat}\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{7}{12} Entropy \left(Sfalse\right)= \frac{7}{12}\*0,985228=0,574716$$

Sehingga Information Gain dari Atribut Fri/Sat adalah

Gain(S, Fri/Sat) = Entropy (S) – Reminder (Fri/Sat)

 = $1- \left(0,404563,+ 0,574716\right)$

 = 0,02072

**Values(Hungry) = True, False**

S = [6+, 6-], |S| = 12

STrue = [5+, 2-], |STrue| = 7

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{5}{7} log\_{2}\left(\frac{5}{7}\right)\right)-\left(\frac{2}{7} log\_{2}\left(\frac{2}{7}\right)\right)$ = 0,863121

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Hungry\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{7}{12} Entropy \left(Strue\right)= \frac{7}{12}\*0,863121=0,503487$$

SFalse = [1+, 4-], |SFalse| = 5

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{1}{5} log\_{2}\left(\frac{1}{5}\right)\right)-\left(\frac{4}{5} log\_{2}\left(\frac{4}{5}\right)\right)$ = 0,721928

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Hungry\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{5}{12} Entropy \left(Sfalse\right)= \frac{5}{12}\*0,721928=0,300803$$

Sehingga Information Gain dari Atribut Hungry adalah

Gain(S, Hungry) = Entropy (S) – Reminder (Hungry)

 = $1- \left(0,503487,+ 0,300803\right)$

 = 0,19571

**Values(Patrons) = None, Some, Full**

S = [6+, 6-], |S| = 12

SNone = [0+, 2-], |SNone| = 2

$Entropy$(SNone) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(SNone) =

$$\sum\_{v ϵ Values \left(Patrons\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{7}{12} Entropy \left(Snone\right)= \frac{7}{12}\*0=0$$

SSome = [4+, 0-], |SSome| = 4

$Entropy$(SSome) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(SSome) =

$$\sum\_{v ϵ Values \left(Patrons\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{7}{12} Entropy \left(Ssome\right)= \frac{7}{12}\*0=0$$

SFull = [2+, 4-], |SFull| = 6

$Entropy$(SFull) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{6} log\_{2}\left(\frac{2}{6}\right)\right)-\left(\frac{4}{6} log\_{2}\left(\frac{4}{6}\right)\right)$ = 0,918296

Reminder(SFull) =

$$\sum\_{v ϵ Values \left(Patrons\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{7}{12} Entropy \left(Sfull\right)= \frac{7}{12}\*0,918296=0,535673$$

Sehingga Information Gain dari Atribut Patrons adalah

Gain(S, Patrons) = Entropy (S) – Reminder (Patrons)

 = $1- \left(0+0+0,535673\right)$

 = 0,464327

**Values(Price) = $, $$, $$$**

S = [6+, 6-], |S| = 12

S$ = [3+, 4-], |S$| = 7

$Entropy$(S$) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{3}{7} log\_{2}\left(\frac{3}{7}\right)\right)-\left(\frac{4}{7} log\_{2}\left(\frac{4}{7}\right)\right)$ = 0,985228

Reminder(S$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{7}{12} Entropy \left(S\$\right)= \frac{7}{12}\*0,985228=0,574716$$

S$$ = [2+, 0-], |S$$| = 2

$Entropy$(S$$) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(S$$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{7}{12} Entropy \left(S\$\$\right)= \frac{7}{12}\*0=0$$

S$$$ = [1+, 2-], |S$$$| = 3

$Entropy$(S$$$) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{1}{3} log\_{2}\left(\frac{1}{3}\right)\right)-\left(\frac{2}{3} log\_{2}\left(\frac{2}{3}\right)\right)$ = 0,918296

Reminder(S$$$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{7}{12} Entropy \left(S\$\$\$\right)= \frac{7}{12}\*0,918296=0,382623$$

Sehingga Information Gain dari Atribut Price adalah

Gain(S, Price) = Entropy (S) – Reminder (Price)

 = $1- \left(0,574716+0+0,382623\right)$

 = 0,04266

**Values(Rain) = True, False**

S = [6+, 6-], |S| = 12

STrue = [2+, 2-], |STrue| = 4

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{4} log\_{2}\left(\frac{2}{4}\right)\right)-\left(\frac{2}{4} log\_{2}\left(\frac{2}{4}\right)\right)$ = 1

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Rain\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{4}{12} Entropy \left(Strue\right)= \frac{4}{12}\*1=0,33333$$

SFalse = [4+, 4-], |SFalse| = 8

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{4}{8} log\_{2}\left(\frac{4}{8}\right)\right)-\left(\frac{4}{8} log\_{2}\left(\frac{4}{8}\right)\right)$ = 1

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Rain\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{8}{12} Entropy \left(Sfalse\right)= \frac{8}{12}\*1=0,66666$$

Sehingga Information Gain dari Atribut Rain adalah

Gain(S, Rain) = Entropy (S) – Reminder (Rain)

 = $1- \left(0,33333+ 0,66666\right)$

 = 0,00001

**Values(Reservation) = True, False**

S = [6+, 6-], |S| = 12

STrue = [3+, 2-], |STrue| = 5

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{3}{5} log\_{2}\left(\frac{3}{5}\right)\right)-\left(\frac{2}{5} log\_{2}\left(\frac{2}{5}\right)\right)$ = 0,970951

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Reservation\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{5}{12} Entropy \left(Strue\right)= \frac{5}{12}\*1=0,404563$$

SFalse = [3+, 4-], |SFalse| = 7

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{3}{7} log\_{2}\left(\frac{3}{7}\right)\right)-\left(\frac{4}{7} log\_{2}\left(\frac{4}{7}\right)\right)$ = 0,985228

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Reservation\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{7}{12} Entropy \left(Sfalse\right)= \frac{7}{12}\*0,985228=0,574716$$

Sehingga Information Gain dari Atribut Reservation adalah

Gain(S, Reservation) = Entropy (S) – Reminder (Reservation)

 = $1- \left(0,404563+ 0,574716\right)$

 = 0,02072

**Values(Type) = Burger, French, Italian, Thai**

S = [6+, 6-], |S| = 12

SBurger = [2+, 2-], |SBurger| = 4

$Entropy$(SBurger) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$1

Reminder(SBurger) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{4}{12} Entropy \left(Sburger\right)= \frac{4}{12}\*1=\frac{4}{12}$$

SFrench = [1+, 1-], |SFrench| = 2

$Entropy$(SFrench) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$1

Reminder(SFrench) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{12} Entropy \left(Sfrench\right)= \frac{2}{12}\*1=\frac{2}{12}$$

SItalian = [1+, 1-], |SItalian| = 2

$Entropy$(SItalian) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $1

Reminder(SItalian) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{12} Entropy \left(Sitalian\right)= \frac{2}{12}\*1=\frac{2}{12}$$

SThai = [2+, 2-], |SThai| = 4

$Entropy$(SThai) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $1

Reminder(SThai) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{4}{12} Entropy \left(Sthai\right)= \frac{4}{12}\*1=\frac{4}{12}$$

Sehingga Information Gain dari Atribut Type adalah

Gain(S, Type) = Entropy (S) – Reminder (Type)

 = $1- \left(\frac{4}{12}+ \frac{2}{12}+\frac{2}{12}+\frac{4}{12}\right)$

 = 0

**Values(Wait Estimate) = 0-10, 10-30, 30-60, >60**

S = [6+, 6-], |S| = 12

S0-10 = [4+, 2-], |S0-10 | = 6

$Entropy$(S0-10) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{4}{6} log\_{2}\left(\frac{4}{6}\right)\right)-\left(\frac{2}{6} log\_{2}\left(\frac{2}{6}\right)\right)=$0,918296

Reminder(S0-10) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{6}{12} Entropy \left(S0-10\right)= \frac{6}{12}\*0,918296=0,459148$$

S10-30= [1+, 1-], |S10-30| = 2

$Entropy$(S10-30) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$1

Reminder(S10-30) =

$$\sum\_{v ϵ Values \left(Wait Estmate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{12} Entropy \left(S10-30\right)= \frac{2}{12}\*1=\frac{2}{12}$$

S30-60= [1+, 1-], |S30-60| = 2

$Entropy$(S30-60) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $1

Reminder(S30-60) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{12} Entropy \left(S30-60\right)= \frac{2}{12}\*1=\frac{2}{12}$$

S>60 = [0+, 2-], |S>60 | = 2

$Entropy$(S>60) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(S>60) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{12} Entropy \left(S>60\right)= \frac{2}{12}\*0=0$$

Sehingga Information Gain dari Atribut Wait Estimate adalah

Gain(S, Wait Estimate) = Entropy (S) – Reminder (Wait Estimate)

 = $1- \left(0,459148+\frac{2}{12}+\frac{2}{12}+0\right)$

 = 0,20752

Tabel Information Gain dari seluruh atribut

|  |  |
| --- | --- |
| Atribut | Information Gain |
| Alternate | 0 |
| Bar | 0 |
| Fri/Sat | 0,02072 |
| Hungry | 0,19571 |
| Patrons | **0,464327** |
| Price | 0,04266 |
| Rain | 0,00001 |
| Reservation | 0,02072 |
| Type | 0 |
| Wait Estimate | 0,20752 |

Berdasarkan tabel Information Gain di atas, diketahui bahwa atribut Patrons memiliki nilai Information Gain terbesar (merupakan atribut yang terbaik dalam mengklasifikasikan data), sehingga **atribut Patrons** akan dijadikan sebagai **ROOT** dari pohon keputusan.

* **Rekursi level 1 iterasi ke 1**

Sampel atribut PatronsSNone = [0+, 2-], |SNone| = 2, karena tiap sampel pada SNone termasuk dalam kelas False dari atribut target maka SNoneakan mengembalikan suatu simpul tunggal ROOT dengan label ‘False’

* **Rekursi level 1 iterasi ke 2**

Sampel atribut HungrySSome = [4+, 0-], |SSome | = 4, karena tiap sampel pada SSometermasuk dalam kelas True dari atribut target maka SSome akan mengembalikan suatu simpul tunggal ROOT dengan label ‘True’

* **Rekursi level 1 iterasi ke 3**

Sampel atribut PatronsSFull = [2+, 4-], |SFull| = 6

Atribut Target = True, False

Kumpulan Atribut = {Alternate, Bar, Fri/Sat, Hungry, Price, Rain, Reservation, Type, Wait Estimate}

Kemudian kita akan mencari nilai Information Gain pada tiap atribut yang ada dengan ruang sampel = SFull

**Values(Alternate) = True, False**

S = SFull = [2+, 4-], |S| = 6

STrue = [2+, 3-], |STrue| = 5

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{5} log\_{2}\left(\frac{2}{5}\right)\right)-\left(\frac{3}{5} log\_{2}\left(\frac{3}{5}\right)\right)$ = 0,970951

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Alternate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{5}{6} Entropy \left(Strue\right)= \frac{5}{6}\*0,970951= 0,809125$$

SFalse = [0+, 1-], |SFalse| = 1

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Alternate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{6} Entropy \left(Sfalse\right)= \frac{1}{6}\*0= 0$$

Sehingga Information Gain dari Atribut Alternate adalah

Gain(S, Alternate) = Entropy (S) – Reminder (Alternate)

 = $1- \left(0,809125+0\right)$

 = 0,19087

**Values(Bar) = True, False**

S = SFull = [2+, 4-], |S| = 6

STrue = [1+, 2-], |STrue| = 3

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi= -\left(\frac{1}{3} log\_{2}\left(\frac{1}{3}\right)\right)-\left(\frac{2}{3} log\_{2}\left(\frac{2}{3}\right)\right)$ = 0,918296

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Bar\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{3}{6} Entropy \left(Strue\right)= \frac{3}{6}\*0,918296= 0,459148$$

SFalse = [1+, 2-], |SFalse| = 3

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi= -\left(\frac{1}{3} log\_{2}\left(\frac{1}{3}\right)\right)-\left(\frac{2}{3} log\_{2}\left(\frac{2}{3}\right)\right)$ = 0,918296

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Bar\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{3}{6} Entropy \left(Sfalse\right)= \frac{3}{6}\*0,918296= 0,459148$$

Sehingga Information Gain dari Atribut Bar adalah

Gain(S, Bar) = Entropy (S) – Reminder (Bar)

 = $1- \left(0,459148+ 0,459148\right)$

 = 0,0817

**Values(Fri/Sat) = True, False**

S = SFull = [2+, 4-], |S| = 6

STrue = [2+, 3-], |STrue| = 5

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{5} log\_{2}\left(\frac{2}{5}\right)\right)-\left(\frac{3}{5} log\_{2}\left(\frac{3}{5}\right)\right)$ = 0,970951

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(\frac{Fri}{Sat}\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{5}{6} Entropy \left(Strue\right)= \frac{5}{6}\*0,970951=0,809125$$

SFalse = [0+, 1-], |SFalse| = 1

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$0

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(\frac{Fri}{Sat}\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{6} Entropy \left(Sfalse\right)= \frac{1}{6}\*0=0$$

Sehingga Information Gain dari Atribut Fri/Sat adalah

Gain(S, Fri/Sat) = Entropy (S) – Reminder (Fri/Sat)

 = $1- \left(0,809125\right)$

 = 0,19087

**Values(Hungry) = True, False**

S = SFull = [2+, 4-], |S| = 6

STrue = [2+, 2-], |STrue| = 4

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi$= 1

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Hungry\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{4}{6} Entropy \left(Strue\right)= \frac{4}{6}\*1=0,66666$$

SFalse = [0+, 2-], |SFalse| = 2

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Hungry\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{6} Entropy \left(Sfalse\right)= \frac{2}{6}\*0=0$$

Sehingga Information Gain dari Atribut Hungry adalah

Gain(S, Hungry) = Entropy (S) – Reminder (Hungry)

 = $1- \left(0,66666+ 0\right)$

 = 0,33334

**Values(Price) = $, $$, $$$**

S = SFull = [2+, 4-], |S| = 6

S$ = [2+, 2-], |S$| = 4

$Entropy$(S$) = $\sum\_{i}^{c}-Pi log\_{2}Pi=1$

Reminder(S$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{4}{6} Entropy \left(S\$\right)= \frac{4}{6}\*1=0,66666$$

S = SFull = [2+, 4-], |S| = 6

S$$ = [0+, 0-], |S$$| = 0

$Entropy$(S$$) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(S$$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{0}{6} Entropy \left(S\$\$\right)= \frac{0}{6}\*0=0$$

S = SFull = [2+, 4-], |S| = 6

S$$$ = [0+, 2-], |S$$$| = 2

$Entropy$(S$$$) = $\sum\_{i}^{c}-Pi log\_{2}Pi$= 0

Reminder(S$$$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{6} Entropy \left(S\$\$\$\right)= \frac{2}{6}\*0=0$$

Sehingga Information Gain dari Atribut Price adalah

Gain(S, Price) = Entropy (S) – Reminder (Price)

 = $1- \left(0,66666\right)$

 = 0,33334

**Values(Rain) = True, False**

S = SFull = [2+, 4-], |S| = 6

STrue = [0+, 1-], |STrue| = 1

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Rain\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{6} Entropy \left(Strue\right)= \frac{1}{6}\*0=0$$

SFalse = [2+, 3-], |SFalse| = 5

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{5} log\_{2}\left(\frac{2}{5}\right)\right)-\left(\frac{3}{5} log\_{2}\left(\frac{3}{5}\right)\right)$ = 0,970951

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Rain\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{5}{6} Entropy \left(Sfalse\right)= \frac{5}{6}\*0,970951=0,809125$$

Sehingga Information Gain dari Atribut Rain adalah

Gain(S, Rain) = Entropy (S) – Reminder (Rain)

 = $1- \left(0,809125\right)$

 = 0,19087

**Values(Reservation) = True, False**

S = SFull = [2+, 4-], |S| = 6

STrue = [0+, 2-], |STrue| = 2

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Reservation\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{6} Entropy \left(Strue\right)= \frac{2}{6}\*0=0$$

SFalse = [2+, 2-], |SFalse| = 4

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{4} log\_{2}\left(\frac{2}{4}\right)\right)-\left(\frac{2}{4} log\_{2}\left(\frac{2}{4}\right)\right)$ = 1

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Reservation\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{4}{6} Entropy \left(Sfalse\right)= \frac{4}{6}\*1=0,66666$$

Sehingga Information Gain dari Atribut Reservation adalah

Gain(S, Reservation) = Entropy (S) – Reminder (Reservation)

 = $1- \left(0+ 0,66666\right)$

 = 0,33334

**Values(Type) = Burger, French, Italian, Thai**

S = SFull = [2+, 4-], |S| = 6

SBurger = [1+, 1-], |SBurger| = 2

$Entropy$(SBurger) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$1

Reminder(SBurger) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{6} Entropy \left(Sburger\right)= \frac{2}{6}\*1=0,333333$$

SFrench = [0+, 1-], |SFrench| = 1

$Entropy$(SFrench) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$0

Reminder(SFrench) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{6} Entropy \left(Sfrench\right)= \frac{1}{6}\*0=0$$

SItalian = [0+, 1-], |SItalian| = 1

$Entropy$(SItalian) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(SItalian) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{6} Entropy \left(Sitalian\right)= \frac{1}{6}\*0=0$$

SThai = [1+, 1-], |SThai| = 2

$Entropy$(SThai) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $1

Reminder(SThai) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{6} Entropy \left(Sthai\right)= \frac{2}{6}\*1=0,333333$$

Sehingga Information Gain dari Atribut Type adalah

Gain(S, Type) = Entropy (S) – Reminder (Type)

 = $1- \left(\frac{2}{6}+ 0+0+\frac{2}{6}\right)$

 = 0,33334

**Values(Wait Estimate) = 0-10, 10-30, 30-60, >60**

S = SFull = [2+, 4-], |S| = 6

S0-10 = [0+, 0-], |S0-10 | = 0

$Entropy$(S0-10) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(S0-10) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)=0$$

S10-30= [1+, 1-], |S10-30| = 2

$Entropy$(S10-30) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$1

Reminder(S10-30) =

$$\sum\_{v ϵ Values \left(Wait Estmate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{6} Entropy \left(S10-30\right)= \frac{2}{6}\*1=0,333333$$

S30-60= [1+, 1-], |S30-60| = 2

$Entropy$(S30-60) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $1

Reminder(S30-60) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{6} Entropy \left(S30-60\right)= \frac{2}{6}\*1=0,333333$$

S>60 = [0+, 2-], |S>60 | = 2

$Entropy$(S>60) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(S>60) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{6} Entropy \left(S>60\right)= \frac{2}{6}\*0=0$$

Sehingga Information Gain dari Atribut Wait Estimate adalah

Gain(S, Wait Estimate) = Entropy (S) – Reminder (Wait Estimate)

 = $1- \left(0+\frac{2}{6}+\frac{2}{6}+0\right)$

 = 0,33334

Tabel Information Gain dari seluruh atribut

|  |  |
| --- | --- |
| Atribut | Information Gain |
| Alternate | 0,19087 |
| Bar | 0,0817 |
| Fri/Sat | 0,19087 |
| Hungry | **0,33334** |
| Price | 0,33334 |
| Rain | 0,19087 |
| Reservation | 0,33334 |
| Type | 0,33334 |
| Wait Estimate | 0,33334 |

Berdasarkan tabel Information Gain di atas, diketahui bahwa atribut **Hungry, Price, Reservation, Type dan Wait Estimate** memiliki nilai Information yang sama besar nilainya, sehingga salah satu atribut akan dipilih menjadi cabang berikutnya yaitu**atribut Hungry.**

* **Rekursi level 2 iterasi ke 1**

Sampel atribut HungrySTrue = [2+, 2-], |STrue| = 4

Atribut Target = True, False

Kumpulan Atribut = {Alternate, Bar, Fri/Sat,Price, Rain, Reservation, Type, Wait Estimate}

Kemudian kita akan mencari nilai Information Gain pada tiap atribut yang ada dengan ruang sampel = STrue

**Values(Alternate) = True, False**

S = STrue = [2+, 2-], |S| = 4

STrue = [2+, 2-], |STrue| = 4

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{4} log\_{2}\left(\frac{2}{4}\right)\right)-\left(\frac{2}{4} log\_{2}\left(\frac{2}{4}\right)\right)$ = 1

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Alternate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{4}{4} Entropy \left(Strue\right)= \frac{4}{4}\*1=1$$

SFalse = [0+, 0-], |SFalse| = 0

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Alternate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)=0$$

Sehingga Information Gain dari Atribut Alternate adalah

Gain(S, Alternate) = Entropy (S) – Reminder (Alternate)

 = $1- \left(1+0\right)$

 = 0

**Values(Bar) = True, False**

S = STrue = [2+, 2-], |S| = 4

STrue = [1+, 1-], |STrue| = 2

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi= -\left(\frac{1}{2} log\_{2}\left(\frac{1}{2}\right)\right)-\left(\frac{1}{2} log\_{2}\left(\frac{1}{2}\right)\right)$ = 1

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Bar\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{4} Entropy \left(Strue\right)= \frac{2}{4}\*1= 0,5$$

SFalse = [1+, 1-], |SFalse| = 2

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi= -\left(\frac{1}{2} log\_{2}\left(\frac{1}{2}\right)\right)-\left(\frac{1}{2} log\_{2}\left(\frac{1}{2}\right)\right)$ = 1

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Bar\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{4} Entropy \left(Sfalse\right)= \frac{2}{4}\*1= 0,5$$

Sehingga Information Gain dari Atribut Bar adalah

Gain(S, Bar) = Entropy (S) – Reminder (Bar)

 = $1- \left(0,5+ 0,5\right)$

 = 0

**Values(Fri/Sat) = True, False**

S = STrue = [2+, 2-], |S| = 4

STrue = [2+, 1-], |STrue| = 3

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{3} log\_{2}\left(\frac{2}{3}\right)\right)-\left(\frac{1}{3} log\_{2}\left(\frac{1}{3}\right)\right)$ = 0,918296

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(\frac{Fri}{Sat}\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{3}{4} Entropy \left(Strue\right)= \frac{3}{4}\*0,918296=0,688722$$

SFalse = [0+, 1-], |SFalse| = 1

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$0

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(\frac{Fri}{Sat}\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{4} Entropy \left(Sfalse\right)= \frac{1}{4}\*0=0$$

Sehingga Information Gain dari Atribut Fri/Sat adalah

Gain(S, Fri/Sat) = Entropy (S) – Reminder (Fri/Sat)

 = $1- \left(0,688722+0\right)$

 = 0,311278

**Values(Price) = $, $$, $$$**

S = STrue = [2+, 2-], |S| = 4

S$ = [2+, 1-], |S$| = 3

$Entropy$(S$) =$\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{3} log\_{2}\left(\frac{2}{3}\right)\right)-\left(\frac{1}{3} log\_{2}\left(\frac{1}{3}\right)\right)=0,918296$

Reminder(S$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{3}{4} Entropy \left(S\$\right)= \frac{3}{4}\*0,918296=0,688722$$

S$$ = [0+, 0-], |S$$| = 0

$Entropy$(S$$) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(S$$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{0}{4} Entropy \left(S\$\$\right)= \frac{0}{4}\*0=0$$

S$$$ = [0+, 1-], |S$$$| = 1

$Entropy$(S$$$) = $\sum\_{i}^{c}-Pi log\_{2}Pi$= 0

Reminder(S$$$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{4} Entropy \left(S\$\$\$\right)= \frac{1}{4}\*0=0$$

Sehingga Information Gain dari Atribut Price adalah

Gain(S, Price) = Entropy (S) – Reminder (Price)

 = $1- \left(0,688722+0+0\right)$

 = 0,311278

**Values(Rain) = True, False**

S = STrue = [0+, 0-], |S| = 0

STrue = [0+, 0-], |STrue| = 0

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Rain\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

SFalse = [2+, 2-], |SFalse| = 4

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{4} log\_{2}\left(\frac{2}{4}\right)\right)-\left(\frac{2}{4} log\_{2}\left(\frac{2}{4}\right)\right)$ = 1

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Rain\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{4}{4} Entropy \left(Sfalse\right)= 1\*1=1$$

Sehingga Information Gain dari Atribut Rain adalah

Gain(S, Rain) = Entropy (S) – Reminder (Rain)

 = $1- \left(0+1\right)$

 = 0

**Values(Reservation) = True, False**

S = STrue = [0+, 1-], |S| = 1

STrue = [0+, 1-], |STrue| = 1

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Reservation\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

SFalse = [2+, 1-], |SFalse| = 3

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{3} log\_{2}\left(\frac{2}{3}\right)\right)-\left(\frac{1}{3} log\_{2}\left(\frac{1}{3}\right)\right)$ =0,918296

0,985228

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Reservation\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{3}{4} Entropy \left(Sfalse\right)= \frac{3}{4}\*0,918295834=0,688721876$$

Sehingga Information Gain dari Atribut Reservation adalah

Gain(S, Reservation) = Entropy (S) – Reminder (Reservation)

 = $1- \left(0+0,688721876\right)$

 = 0,3311278

**Values(Type) = Burger, French, Italian, Thai**

S = STrue = [0+, 0-], |S| = 0

SBurger = [1+, 0-], |SBurger| = 0

$Entropy$(SBurger) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$0

Reminder(SBurger) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

SFrench = [0+, 0-], |SFrench| = 0

$Entropy$(SFrench) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$0

Reminder(SFrench) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

SItalian = [0+, 1-], |SItalian| = 1

$Entropy$(SItalian) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(SItalian) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{6} Entropy \left(Sitalian\right)= \frac{1}{6}\*0=0$$

SThai = [1+, 1-], |SThai| = 2

$Entropy$(SThai) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $1

Reminder(SThai) =

$$\sum\_{v ϵ Values \left(Type\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{4} Entropy \left(Sthai\right)= \frac{2}{4}\*1=0,5$$

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Sehingga Information Gain dari Atribut Type adalah

Gain(S, Type) = Entropy (S) – Reminder (Type)

 = $1- \left(0.5\right)$

 = 0,5

**Values(Wait Estimate) = 0-10, 10-30, 30-60, >60**

S = STrue = [0+, 0-], |S| = 0

S0-10 = [0+, 0-], |S0-10 | = 0

$Entropy$(S0-10) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(S0-10) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)=0$$

S10-30= [1+, 1-], |S10-30| = 2

$Entropy$(S10-30) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$1

Reminder(S10-30) =

$$\sum\_{v ϵ Values \left(Wait Estmate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{4} Entropy \left(S10-30\right)= \frac{2}{4}\*1=0,5$$

S30-60= [1+, 1-], |S30-60| = 2

$Entropy$(S30-60) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $1

Reminder(S30-60) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{4} Entropy \left(S30-60\right)= \frac{2}{4}\*1=0,5$$

S>60 = [0+, 0-], |S>60 | = 0

$Entropy$(S>60) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(S>60) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

Sehingga Information Gain dari Atribut Wait Estimate adalah

Gain(S, Wait Estimate) = Entropy (S) – Reminder (Wait Estimate)

 = $1- \left(0+0.5+0.5+0\right)$

 = 0

Tabel Information Gain dari seluruh atribut

|  |  |
| --- | --- |
| Atribut | Information Gain |
| Alternate | 0 |
| Bar | 0 |
| Fri/Sat | 0,311278 |
| Price | 0,311278 |
| Rain | 0 |
| Reservation | 0,3311278 |
| Type | **0,5** |
| Wait Estimate | 0 |

Berdasarkan tabel Information Gain di atas, diketahui bahwa **atribut Type** memiliki nilai Information Gain terbesar (merupakan atribut yang terbaik dalam mengklasifikasikan data), sehingga **atribut Type**akan dijadikan cabang berikutnya.

* **Rekursi level 3 iterasi ke 1**

Sampel atribut Type SFrench = [0+, 0-], |SFrench| = 0

Karena SFrench tidak memiliki sampel, maka SFrench akan mengembalikan nilai suatu simpul tunggal dengan label yang paling dominan pada Sampel HungryTrue yaitu ‘True’

Sampel atribut Type SItalian = [0+, 1-], |SItalian| = 1, karena tiap sampel pada SItalian termasuk dalam kelas False dari atribut target maka SItalian akan mengembalikan suatu simpul tunggal pada pohon keputusan dengan label ‘False’

* **Rekursi level 3 iterasi ke 2**

Sampel atribut Type SThai = [1+, 1-], |SThai| = 2Atribut Target = True, False

Kumpulan Atribut = {Alternate, Bar, Fri/Sat,Price, Rain, Reservation, Wait Estimate}

Kemudian kita akan mencari nilai Information Gain pada tiap atribut yang ada dengan ruang sampel = SThai

**Values(Alternate) = True, False**

S = SThai = [1+, 1-], |S| = 2

STrue = [1+, 1-], |STrue| = 2

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=-\left(\frac{2}{4} log\_{2}\left(\frac{2}{4}\right)\right)-\left(\frac{2}{4} log\_{2}\left(\frac{2}{4}\right)\right)$ = 1

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Alternate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{2} Entropy \left(Strue\right)= \frac{1}{1}\*1=1$$

SFalse = [0+, 0-], |SFalse| = 0

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Alternate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)=0$$

Sehingga Information Gain dari Atribut Alternate adalah

Gain(S, Alternate) = Entropy (S) – Reminder (Alternate)

 = $1- \left(1+0\right)$

 = 0

**Values(Bar) = True, False**

S = SThai = [1+, 1-], |S| = 2

STrue = [0+, 0-], |STrue| = 0

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi= 0$

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Bar\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

SFalse = [1+, 1-], |SFalse| = 2

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi= -\left(\frac{1}{2} log\_{2}\left(\frac{1}{2}\right)\right)-\left(\frac{1}{2} log\_{2}\left(\frac{1}{2}\right)\right)$ = 1

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Bar\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{2}{2} Entropy \left(Sfalse\right)= 1$$

Sehingga Information Gain dari Atribut Bar adalah

Gain(S, Bar) = Entropy (S) – Reminder (Bar)

 = $1- \left(0+ 1\right)$

 = 0

**Values(Fri/Sat) = True, False**

S = SThai = [1+, 1-], |S| = 2

STrue = [1+, 0-], |STrue| = 1

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(\frac{Fri}{Sat}\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

SFalse = [0+, 1-], |SFalse| = 1

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$0

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(\frac{Fri}{Sat}\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{4} Entropy \left(Sfalse\right)= \frac{1}{4}\*0=0$$

Sehingga Information Gain dari Atribut Fri/Sat adalah

Gain(S, Fri/Sat) = Entropy (S) – Reminder (Fri/Sat)

 = $1- \left(0+0\right)$

 = 1

**Values(Price) = $, $$, $$$**

S = SThai = [1+, 1-], |S| = 2

S$ = [1+, 1-], |S$| = 2

$Entropy$(S$) =$\sum\_{i}^{c}-Pi log\_{2}Pi=1$

Reminder(S$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 1$$

S = SThai = [1+, 1-], |S| = 2

S$$ = [0+, 0-], |S$$| = 0

$Entropy$(S$$) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(S$$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{0}{2} Entropy \left(S\$\$\right)= \frac{0}{2}\*0=0$$

S = SThai = [1+, 1-], |S| = 2

S$$$ = [0+, 1-], |S$$$| = 1

$Entropy$(S$$$) = $\sum\_{i}^{c}-Pi log\_{2}Pi$= 0

Reminder(S$$$) =

$$\sum\_{v ϵ Values \left(Price\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{2} Entropy \left(S\$\$\$\right)= \frac{1}{2}\*0=0$$

Sehingga Information Gain dari Atribut Price adalah

Gain(S, Price) = Entropy (S) – Reminder (Price)

 = $1- \left(1+0+0\right)$

 = 0

**Values(Rain) = True, False**

S = SThai = [1+, 1-], |S| = 2

STrue = [0+, 0-], |STrue| = 0

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Rain\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

SFalse = [1+, 1-], |SFalse| = 2

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi$= 1

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Rain\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{2} Entropy \left(Sfalse\right)= 0.5$$

Sehingga Information Gain dari Atribut Rain adalah

Gain(S, Rain) = Entropy (S) – Reminder (Rain)

 = $1- \left(0+0.5\right)$

 = 0.5

**Values(Reservation) = True, False**

S = SThai = [1+, 1-], |S| = 2

STrue = [0+, 0-], |STrue| = 0

$Entropy$(STrue) = $\sum\_{i}^{c}-Pi log\_{2}Pi=0$

Reminder(STrue) =

$$\sum\_{v ϵ Values \left(Reservation\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

SFalse = [1+, 1-], |SFalse| = 2

$Entropy$(SFalse) = $\sum\_{i}^{c}-Pi log\_{2}Pi=1$

0,985228

Reminder(SFalse) =

$$\sum\_{v ϵ Values \left(Reservation\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= \frac{1}{2} Entropy \left(Sfalse\right)= 0,5$$

Sehingga Information Gain dari Atribut Reservation adalah

Gain(S, Reservation) = Entropy (S) – Reminder (Reservation)

 = $1- \left(0+0,5\right)$

 = 0,5

**Values(Wait Estimate) = 0-10, 10-30, 30-60, >60**

S = SThai = [1+, 1-], |S| = 2

S0-10 = [0+, 0-], |S0-10 | = 0

$Entropy$(S0-10) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(S0-10) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)=0$$

S10-30= [1+, 0-], |S10-30| = 1

$Entropy$(S10-30) = $\sum\_{i}^{c}-Pi log\_{2}Pi=$ 0

Reminder(S10-30) =

$$\sum\_{v ϵ Values \left(Wait Estmate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

S30-60= [0+, 1-], |S30-60| = 1

$Entropy$(S30-60) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(S30-60) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

S>60 = [0+, 0-], |S>60 | = 0

$Entropy$(S>60) = $\sum\_{i}^{c}-Pi log\_{2}Pi= $0

Reminder(S>60) =

$$\sum\_{v ϵ Values \left(Wait Estimate\right)}^{}\frac{\left|Sv\right|}{\left|S\right|} Entropy\left(Sv\right)= 0$$

Sehingga Information Gain dari Atribut Wait Estimate adalah

Gain(S, Wait Estimate) = Entropy (S) – Reminder (Wait Estimate)

 = $1- \left(0+0+0+0\right)$

 = 1

Tabel Information Gain dari seluruh atribut

|  |  |
| --- | --- |
| Atribut | Information Gain |
| Alternate | 0 |
| Bar | 0 |
| Fri/Sat | **1** |
| Price | 0 |
| Rain | 0,5 |
| Reservation | 0,5 |
| Wait Estimate | 1 |

Berdasarkan tabel Information Gain di atas, diketahui bahwa **atribut Fri/Sat** memiliki nilai Information Gain terbesar (merupakan atribut yang terbaik dalam mengklasifikasikan data), sehingga **atributFri/Sat** akan dijadikan cabang berikutnya.

* **Rekursi level 4 iterasi ke 1**

Sampel atribut Fri/Sat STrue = [1+, 0-], |STrue| = 1, karena tiap sampel pada STrue termasuk dalam kelas True dari atribut target maka STrueakan mengembalikan suatu simpul tunggal pada pohon keputusan dengan label ‘True’

* **Rekursi level 4 iterasi ke 2**

Sampel atribut Fri/Sat SFalse = [0+, 1-], |SFalse| = 1, karena tiap sampel pada SFalse termasuk dalam kelas False dari atribut target maka SFalseakan mengembalikan suatu simpul tunggal pada pohon keputusan dengan label ‘False’

* **Rekursi level 3 iterasi ke 4**

Sample atribut Type SBurger = [1+, 0-], |SBurger| = 1, karena tiap sampel pada SBurger termasuk dalam kelas True dari atribut target maka SBurger akan mengembalikan suatu simpul tunggal pada pohon keputusan dengan label ‘True’

* **Rekursi level 2 iterasi ke 4**

Sampel atribut Hungry SFalse = [0+, 2-], |SFalse| = 2, karena tiap sampel pada SFalse termasuk dalam kelas True dari atribut target maka SFalseakan mengembalikan suatu simpul tunggal pada pohon keputusan dengan label ‘False’