**GINI working**:

P(yes)\*P(job=no, status=single, income=60|approve=yes)=

8/12\*3/9\*4/5\*0.00360255296=0.00064045385

P(no)\*P(job=no, status=single, income=60|approve=no)=

4/12\*4/5\*2/5\*0.05174060994=**0.00551899839** **<- higher**

*YesIncome=60***:**1/sqrt(2pi\*18.1920)exp(-(60-(745/7))^2/(2\*330.95238))= 0.0036

*NoIncome=60*:1/sqrt(2pi\*25.16611)exp(-(60-(83+1/3))^2/(2\*633.3333))= 0.05174

P(yes)\*P(job=yes, status=married, income=100|approve=yes)=

8/12\*6/9\*4/5\*0.08787253995=**0.03124356976 <- higher**

P(no)\*P(job=yes, status=married, income=100|approve=no)=

4/12\*1/5\*2/5\*0.06386485953=0.00170306292

*YesIncome=100***:**1/sqrt(2pi\*18.192)exp(-(100-(745/7))^2/(2\*330.95238))= 0.087872

*NoIncome=100*:1/sqrt(2pi\*25.16611)exp(-(100-(83+1/3))^2/(2\*633.33))= 0.0638648

P(yes)\*P(job=yes, status=single, income=90|approve=yes)=

8/12\*6/9\*4/5\*0.06221=**0.02212 <- higher**

P(no)\*P(job=yes, status=single, income=90|approve=no)=

4/12\*1/5\*2/5\*0.07678=0.002047539

*YesIncome=90***:**1/sqrt(2pi\*18.19209)exp(-(90-(745/7))^2/(2\*330.95238))= 0.0622129

*NoIncome=90*:1/sqrt(2pi\*25.16611)exp(-(90-(83+1/3))^2/(2\*633.3333))= 0.0767827

P(yes)\*P(job=no, status=divorced, income=95|approve=yes)=

8/12\*3/9\*2/5\*0.07678=0.0068252

P(no)\*P(job=no, status=divorced, income=95|approve=no)=

4/12\*4/5\*2/5\*0.071422=**0.0076183** **<- higher**

*YesIncome=95***:**1/sqrt(2pi\*18.19209)exp(-(95-(745/7))^2/(2\*330.95238))= 0.0767839

*NoIncome=95*:1/sqrt(2pi\*25.16611)exp(-(95-(83+1/3))^2/(2\*633.3333))= 0.07142238

P(yes)\*P(job=no, status=married, income=85|approve=yes)=

8/12\*3/9\*4/5\*0.046739=0.008309

P(no)\*P(job=no, status=married, income=85|approve=no)=

4/12\*4/5\*2/5\*0.07935=**0.008464** **<- higher**

*YesIncome=85***:**1/sqrt(2pi\*18.19209)exp(-(85-(745/7))^2/(2\*330.95238))= 0.04673962

*NoIncome=85*:1/sqrt(2pi\*25.16611)exp(-(85-(83+1/3))^2/(2\*633.3333))= 0.07935048

**Working for Naïve Bayes (using Laplacian correction):**

P(yes) = 8/12 P(no) = 4/12

for feature *Permanent Job and* *Marital Status*

approve? = no: 3 examples

* Permanent Job = yes: 0 examples
  + P(Permanent Job = yes | approve? = no) = 1/5
* Permanent Job = no: 3 examples
  + P(Permanent Job = no | approve? = no) = 4/5
* Marital Status = single: 1 example
  + P(Permanent Job = single | approve? = no) = 2/5
* Marital Status = married: 1 example
  + P(Permanent Job = married | approve? = no) = 2/5
* Marital Status = divorced: 1 example
  + P(Permanent Job = divorced | approve? = no) = 2/5

approve? = yes: 7 examples

* Permanent Job = yes: 5 examples
  + P(Permanent Job = yes | approve? = yes) = 6/9
* Permanent Job = no: 2 examples
  + P(Permanent Job = no | approve? = yes) = 3/9
* Marital Status = single: 3 examples
  + P(Permanent Job = single | approve? = yes) = 4/5
* Marital Status = married: 3 examples
  + P(Permanent Job = married | approve? = yes) = 4/5
* Marital Status = divorced: 1 example
  + P(Permanent Job = divorced | approve? = yes) = 2/5

*Annual Income:*

Data of class no:

* Miu: (80+60+110)/3 = 250/3
* Sigma^2: ((80-(250/3))^2+(60-(250/3))^2+(110-(250/3))^2)/2 = 633.333333333
* Sigma: 25.1661147842

P(number | approve? = no) = 1/sqrt(2pi\*25.16611)exp(-(number-(83+1/3))^2/(2\*633.333))

Data of class yes:

* Miu: (130+100+90+120+85+95+125)/7 = 745/7
* Sigma^2: ((130-(745/7))^2+(100-(745/7))^2+(90-(745/7))^2+(120-(745/7))^2+(85-(745/7))^2+(95-(745/7))^2+(125-(745/7))^2)/6 = 330.952380952
* Sigma: 18.1920966618

P(number | approve? = yes) = 1/sqrt(2pi\*18.19209)exp(-(number-(745/7))^2/(2\*330.95238))

**Working for Decision Tree:** IV(Permanent Job) = = 1

IV(Marital Status) = -4/10\*lb(4/10)-4/10\*lb(4/10)-2/10\*lb(2/10) = 1.52192809489

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 60 | | 80 | | 85 | | 90 | | 95 | | 100 | | 110 | | 120 | | 125 | 130 |
| 70 | | 82.5 | | 87.5 | | 92.5 | | 97.5 | | 105 | | 115 | | 122.5 | 127.5 |

*70: IV(Annual Income | 70)* = -1/10\*log(2, 1/10)-9/10\*log(2, 9/10) = 0.4689956

Gain({1-10}, Annual Income, 70) = Ent({1-10}) - … = 0.88129 (-0/1\*log(2,0/1)-1/1\*log(2,1/1)) (-7/9\*log(2,7/9)-2/9\*log(2,2/9)) = 0.88129-1/10\*0-9/10\*0.7642=0.19351

Gain\_ratio({1-10}, Annual Income)= 0.19351/0.4689956=0.4126

*82.5: IV(Annual Income | 82.5)* = -2/10\*log(2, 2/10)-8/10\*log(2, 8/10) = 0.721928

Gain({1-10}, Annual Income, 82.5) = Ent({1-10}) - … = 0.88129 (-0/2\*log(2,0/2)-2/2\*log(2,2/2)) (-7/8\*log(2,7/8)-1/8\*log(2,1/8)) = 0.88129-2/10\*0-8/10\*0.54356=0.446442

Gain\_ratio({1-10}, Annual Income)= 0.446442/0.721928=**0.6184 <- highest**

*87.5: IV(Annual Income | 87.5)* = -3/10\*log(2, 3/10)-7/10\*log(2, 7/10) = 0.88129

Gain({1-10}, Annual Income, 87.5) = Ent({1-10}) - … = 0.88129 (-1/3\*log(2,1/3)-2/3\*log(2,2/3)) (-6/7\*log(2,6/7)-1/7\*log(2,1/7)) = 0.88129-3/10\*0.9183-7/10\*0.59167=0.191631

Gain\_ratio({1-10}, Annual Income)= 0.191631/0.88129=0.191631/0.88129

**…**

Gain({1-10}, Permanent Job) = Ent({1-10}) - … = (-7/10\*log(2,7/10)- 3/10\*log(2,3/10)) (- 5/5\*log(2,5/5)- 0/5\*log(2,0/5)) (-2/5\*log(2,2/5)- 3/5\*log(2,3/5)) = 0.88129-0-5/10\*(0.9709)=0.39584

Gain({1-10}, Marital Status) = Ent({1-10}) - … = 0.88129 (-3/4\*log(2,3/4)- 1/4\*log(2,1/4)) (-3/4\*log(2,3/4)-1/4\*log(2,1/4)) (-1/2\*log(2,1/2)-1/2\*log(2,1/2)) = 0.88129-4/10\*0.811278-4/10\*0.811278-2/10\*1=0.0322676

Gain({1-10}, Annual Income) = Ent({1-10}) - … = 0.88129 (-4/5\*log(2,4/5)-1/5\*log(2,1/5)) (-3/5\*log(2,3/5)-2/5\*log(2,2/5)) = 0.88129-5/10\*0.72192809488-5/10\*0.97095059445=0.03485065533

Gain\_ratio({1-10}, Permanent Job)= 0.39584/1=0.39584

Gain\_ratio({1-10}, Marital Status)= 0.0322676/1.52192809489=0.02120179009

Gain\_ratio({1-10}, Annual Income)= 0.03485065533/1=0.03485065533

|  |  |  |  |
| --- | --- | --- | --- |
| Permanent Job | Marital Status | Annual Income | Approved? |
| Yes | Single | *130K* | Yes |
| No | Single | *100K* | Yes |
| Yes | Divorced | *90K* | Yes |
| Yes | Married | *120K* | Yes |
| Yes | Single | *85K* | Yes |
| No | Divorced | *110K* | No |
| Yes | Married | *95K* | Yes |
| No | Married | *125K* | Yes |

IV(Permanent Job) = -3/8\*log(2,3/8)-5/8\*log(2,5/8) = 0.954434

IV(Marital Status) = -3/8\*log(2,3/8)-3/8\*log(2,3/8)-2/8\*log(2,2/8) = 1.561278

Gain({1-8}, Permanent Job) = Ent({1-8}) - … = 0.97095 (-2/3\*log(2,2/3)-1/3\*log(2,1/3)) (-0/5\*log(2,0/5)-5/5\*log(2,5/5)) = 0.54356444-3/8\*0.9182958-0=0.199203515

Gain({1-8}, Marital Status) = Ent({1-8}) - … = (-1/8\*log(2,1/8)- 7/8\*log(2,7/8)) (-3/3\*log(2,3/3)- 0/3\*log(2,0/3)) (-1/2\*log(2,1/2)-1/2\*log(2,1/2)) (-0/3\*log(2,0/3)-3/3\*log(2,3/3)) = 0.54356444-0-2/8\*1-0=0.29356444 **<- highest**

Gain\_ratio({1-8}, Permanent Job)= 0.29356444 /0.954434=**0.30758** **<- highest**

Gain\_ratio({1-8}, Marital Status)= 0.199203515/1.561278=0.12759