**135356**

DECISION TREES

Risk:Low=4/10

Medium=3/10

High=3/10

Calculating the OG entropy

Class:

E(S)= E(low)+E(Medium)+E(High)

-P(I) log2 P(I)

(-4/10 log2(4/10)) + (-3/10 log2(3/10)) + (-3/10 log2(3/10))

=0.5287+0.5210+0.5210=1.5709

1. Debt

Low= 5/10 low risk=4/5

medium risk=1/5

Medium=2/10 medium risk=2/2

High=3/10 high risk=3/3

E (S low) = 5/10((-4/5 log2(4/5)) + (-1/5 log2(1/5))) = 0.36095

E (S medium) = (-2/2 log2(2/2)) = 0

E (S high) = (-3/3 log2(3/3)) = 0

E (S debt) = 0.36095+0+0= 0.36095

Information Gain=OG Entropy-Debt

Debt Information Gain = 1.5709-0.36095=1.20995

1. Collateral

None= 5/10 low risk=3/5

medium risk=1/5

high risk=1/5

Adequate= 5/10 low risk=1/5

medium risk=2/5

high risk=2/5

E(S none) = 5/10((-3/5 log2(3/5)) + (-1/5 log2(1/5)) + (-1/5 log2(1/5))) =0.6854

E(S adequate) = 5/10((-1/5 log2(1/5)) + (-2/5 log2(2/5)) + (-2/5 log2(2/5))) =0.76095

E(S collateral) = 0.6854+0.76095= 1.4464

Collateral IG = 1.5709-1.4464= 0.1245

1. Credit History

Bad= 2/10->low risk=2/2

Unknown=6/10 low risk=2/6

medium risk=3/6

high risk=1/6

Good=2/10->high risk=2/2

E(S bad) = (-2/2 log2(2/2)) = 0

E(S known) = 6/10((-2/6 log2(2/6)) + (-3/6 log2(3/6)) + (-1/6 log2(1/6))) = 0.8755

E(S good) = (-2/2 log2(2/2)) = 0

E(S credit) = 0+0.8755+0= 0.8755

Credit IG=1.5709-0.8755= 0.6954

1. Income shs‘000’

0-15= 3/10 low risk=2/3

med risk=1/3

15-35= 5/10 low risk=2/5

med risk=2/5

high risk=1/5

Over 35= 2/10 high risk=2/2

E(S0-15) = 3/10((-2/3 log2(2/3)) + (-1/3 log2(1/3))) = 0.27549

E(S15-35) = 5/10((-2/5 log2(2/5)) + (-2/5 log2(2/5)) + (-1/5 log2(1/5))) = 0.76095

E(Sover35) = (-2/2 log2(2/2)) = 0

E(S credit) = 0.275449+0.76095+0= 1.03645

Income IG =1.5709-1.03645= 0.53445