HW1 Decision Tree Algorithm

Haoran Lou hxl180012

Result:

1. data set 1

Pruned tree for L, K = (85,45), just one example because of 20 trees will be toooo long.

XO = 1:

| XI = 1:

| | XT = 1:

| | | XS = 1:

| | | | XL = 1:

| | | | | XH = 1: 0

| | | | | XH=0:

| | | | | | XD = 1:

| | | | | | | XB = 1: 0

| | | | | | | XB=0: 1

| | | | | | XD=0:

| | | | | | | XQ = 1:

| | | | | | | | XB = 1: 1

| | | | | | | | XB=0: 0

| | | | | | | XQ=0: 0

| | | | XL=0:

| | | | | XD = 1:

| | | | | | XG = 1: 1

| | | | | | XG=0: 0

| | | | | XD=0:

| | | | | | XU = 1:

| | | | | | | XF = 1:

| | | | | | | | XE = 1:

| | | | | | | | | XC = 1: 0

| | | | | | | | | XC=0: 1

| | | | | | | | XE=0: 1

| | | | | | | XF=0:

| | | | | | | | XG = 1: 0

| | | | | | | | XG=0:

| | | | | | | | | XH = 1: 1

| | | | | | | | | XH=0: 0

| | | | | | XU=0: 1

| | | XS=0:

| | | | XQ = 1:

| | | | | XM = 1: 0

| | | | | XM=0:

| | | | | | XN = 1:

| | | | | | | XP = 1: 1

| | | | | | | XP=0:

| | | | | | | | XB = 1: 1

| | | | | | | | XB=0:

| | | | | | | | | XF = 1: 1

| | | | | | | | | XF=0: 0

| | | | | | XN=0:

| | | | | | | XU = 1: 0

| | | | | | | XU=0: 1

| | | | XQ=0:

| | | | | XK = 1:

| | | | | | XD = 1: 0

| | | | | | XD=0:

| | | | | | | XF = 1: 1

| | | | | | | XF=0: 0

| | | | | XK=0:

| | | | | | XC = 1: 1

| | | | | | XC=0:

| | | | | | | XR = 1:

| | | | | | | | XB = 1: 0

| | | | | | | | XB=0:

| | | | | | | | | XD = 1: 1

| | | | | | | | | XD=0: 0

| | | | | | | XR=0:

| | | | | | | | XH = 1: 1

| | | | | | | | XH=0:

| | | | | | | | | XE = 1: 1

| | | | | | | | | XE=0: 0

| | XT=0:

| | | XH = 1:

| | | | XJ = 1:

| | | | | XS = 1:

| | | | | | XG = 1:

| | | | | | | XC = 1:

| | | | | | | | XD = 1: 0

| | | | | | | | XD=0: 1

| | | | | | | XC=0: 1

| | | | | | XG=0:

| | | | | | | XB = 1:

| | | | | | | | XD = 1:

| | | | | | | | | XE = 1: 1

| | | | | | | | | XE=0: 0

| | | | | | | | XD=0: 1

| | | | | | | XB=0: 0

| | | | | XS=0: 1

| | | | XJ=0:

| | | | | XC = 1:

| | | | | | XM = 1:

| | | | | | | XF = 1: 1

| | | | | | | XF=0:

| | | | | | | | XR = 1: 0

| | | | | | | | XR=0: 1

| | | | | | XM=0: 0

| | | | | XC=0:

| | | | | | XN = 1:

| | | | | | | XF = 1: 0

| | | | | | | XF=0:

| | | | | | | | XG = 1: 0

| | | | | | | | XG=0: 1

| | | | | | XN=0: 1

| | | XH=0:

| | | | XP = 1:

| | | | | XS = 1: 1

| | | | | XS=0:

| | | | | | XD = 1:

| | | | | | | XM = 1: 1

| | | | | | | XM=0:

| | | | | | | | XC = 1: 0

| | | | | | | | XC=0: 1

| | | | | | XD=0:

| | | | | | | XC = 1: 0

| | | | | | | XC=0:

| | | | | | | | XJ = 1:

| | | | | | | | | XG = 1: 0

| | | | | | | | | XG=0:

| | | | | | | | | | XB = 1: 1

| | | | | | | | | | XB=0: 0

| | | | | | | | XJ=0:

| | | | | | | | | XN = 1: 1

| | | | | | | | | XN=0: 0

| | | | XP=0:

| | | | | XF = 1:

| | | | | | XQ = 1:

| | | | | | | XK = 1: 1

| | | | | | | XK=0: 0

| | | | | | XQ=0:

| | | | | | | XK = 1:

| | | | | | | | XC = 1: 1

| | | | | | | | XC=0: 0

| | | | | | | XK=0: 1

| | | | | XF=0: 0

| XI=0:

| | XM = 1:

| | | XQ = 1: 0

| | | XQ=0:

| | | | XF = 1: 0

| | | | XF=0:

| | | | | XL = 1:

| | | | | | XC = 1: 1

| | | | | | XC=0:

| | | | | | | XB = 1:

| | | | | | | | XP = 1: 1

| | | | | | | | XP=0: 0

| | | | | | | XB=0: 0

| | | | | XL=0:

| | | | | | XC = 1:

| | | | | | | XH = 1:

| | | | | | | | XU = 1: 1

| | | | | | | | XU=0:

| | | | | | | | | XB = 1: 0

| | | | | | | | | XB=0:

| | | | | | | | | | XD = 1: 0

| | | | | | | | | | XD=0: 1

| | | | | | | XH=0: 1

| | | | | | XC=0: 1

| | XM=0:

| | | XQ = 1:

| | | | XJ = 1:

| | | | | XL = 1: 1

| | | | | XL=0:

| | | | | | XH = 1:

| | | | | | | XK = 1: 1

| | | | | | | XK=0:

| | | | | | | | XU = 1: 0

| | | | | | | | XU=0: 1

| | | | | | XH=0: 0

| | | | XJ=0:

| | | | | XN = 1: 0

| | | | | XN=0:

| | | | | | XP = 1:

| | | | | | | XB = 1: 0

| | | | | | | XB=0:

| | | | | | | | XF = 1: 1

| | | | | | | | XF=0: 0

| | | | | | XP=0: 1

| | | XQ=0:

| | | | XF = 1: 0

| | | | XF=0:

| | | | | XH = 1: 1

| | | | | XH=0:

| | | | | | XB = 1:

| | | | | | | XC = 1: 0

| | | | | | | XC=0: 1

| | | | | | XB=0: 0

XO=0:

| XM = 1:

| | XB = 1:

| | | XI = 1:

| | | | XC = 1: 0

| | | | XC=0:

| | | | | XK = 1: 0

| | | | | XK=0:

| | | | | | XP = 1:

| | | | | | | XS = 1: 0

| | | | | | | XS=0:

| | | | | | | | XG = 1:

| | | | | | | | | XF = 1: 1

| | | | | | | | | XF=0: 0

| | | | | | | | XG=0: 1

| | | | | | XP=0: 1

| | | XI=0: 0

| | XB=0:

| | | XF = 1:

| | | | XS = 1:

| | | | | XR = 1:

| | | | | | XI = 1:

| | | | | | | XG = 1: 1

| | | | | | | XG=0:

| | | | | | | | XQ = 1: 0

| | | | | | | | XQ=0:

| | | | | | | | | XJ = 1: 1

| | | | | | | | | XJ=0:

| | | | | | | | | | XD = 1: 1

| | | | | | | | | | XD=0: 0

| | | | | | XI=0:

| | | | | | | XU = 1: 0

| | | | | | | XU=0: 1

| | | | | XR=0:

| | | | | | XD = 1:

| | | | | | | XK = 1: 1

| | | | | | | XK=0:

| | | | | | | | XE = 1:

| | | | | | | | | XG = 1: 0

| | | | | | | | | XG=0: 1

| | | | | | | | XE=0: 0

| | | | | | XD=0: 1

| | | | XS=0:

| | | | | XQ = 1:

| | | | | | XI = 1:

| | | | | | | XL = 1:

| | | | | | | | XD = 1: 1

| | | | | | | | XD=0: 0

| | | | | | | XL=0: 1

| | | | | | XI=0: 1

| | | | | XQ=0:

| | | | | | XI = 1:

| | | | | | | XP = 1: 1

| | | | | | | XP=0: 0

| | | | | | XI=0: 0

| | | XF=0:

| | | | XG = 1:

| | | | | XD = 1:

| | | | | | XC = 1: 0

| | | | | | XC=0:

| | | | | | | XP = 1: 0

| | | | | | | XP=0:

| | | | | | | | XS = 1: 1

| | | | | | | | XS=0: 0

| | | | | XD=0: 1

| | | | XG=0: 0

| XM=0:

| | XF = 1: 0

| | XF=0:

| | | XB = 1:

| | | | XD = 1:

| | | | | XI = 1:

| | | | | | XG = 1: 0

| | | | | | XG=0: 1

| | | | | XI=0: 0

| | | | XD=0: 0

| | | XB=0:

| | | | XG = 1:

| | | | | XD = 1:

| | | | | | XP = 1: 0

| | | | | | XP=0:

| | | | | | | XE = 1: 1

| | | | | | | XE=0: 0

| | | | | XD=0:

| | | | | | XS = 1:

| | | | | | | XC = 1:

| | | | | | | | XH = 1: 1

| | | | | | | | XH=0: 0

| | | | | | | XC=0: 1

| | | | | | XS=0: 0

| | | | XG=0: 0

XO = 1:

| XI = 1:

| | XT = 1:

| | | XS = 1:

| | | | XL = 1:

| | | | | XH = 1: 0

| | | | | XH=0:

| | | | | | XD = 1:

| | | | | | | XB = 1: 0

| | | | | | | XB=0: 1

| | | | | | XD=0:

| | | | | | | XQ = 1:

| | | | | | | | XB = 1: 1

| | | | | | | | XB=0: 0

| | | | | | | XQ=0: 0

| | | | XL=0:

| | | | | XD = 1:

| | | | | | XG = 1: 1

| | | | | | XG=0: 0

| | | | | XD=0:

| | | | | | XU = 1:

| | | | | | | XF = 1:

| | | | | | | | XE = 1:

| | | | | | | | | XC = 1: 0

| | | | | | | | | XC=0: 1

| | | | | | | | XE=0: 1

| | | | | | | XF=0:

| | | | | | | | XG = 1: 0

| | | | | | | | XG=0:

| | | | | | | | | XH = 1: 1

| | | | | | | | | XH=0: 0

| | | | | | XU=0: 1

| | | XS=0:

| | | | XQ = 1:

| | | | | XM = 1: 0

| | | | | XM=0:

| | | | | | XN = 1:

| | | | | | | XP = 1: 1

| | | | | | | XP=0:

| | | | | | | | XB = 1: 1

| | | | | | | | XB=0:

| | | | | | | | | XF = 1: 1

| | | | | | | | | XF=0: 0

| | | | | | XN=0:

| | | | | | | XU = 1: 0

| | | | | | | XU=0: 1

| | | | XQ=0:

| | | | | XK = 1:

| | | | | | XD = 1: 0

| | | | | | XD=0:

| | | | | | | XF = 1: 1

| | | | | | | XF=0: 0

| | | | | XK=0:

| | | | | | XC = 1: 1

| | | | | | XC=0:

| | | | | | | XR = 1:

| | | | | | | | XB = 1: 0

| | | | | | | | XB=0:

| | | | | | | | | XD = 1: 1

| | | | | | | | | XD=0: 0

| | | | | | | XR=0:

| | | | | | | | XH = 1: 1

| | | | | | | | XH=0:

| | | | | | | | | XE = 1: 1

| | | | | | | | | XE=0: 0

| | XT=0:

| | | XH = 1:

| | | | XJ = 1:

| | | | | XS = 1:

| | | | | | XG = 1:

| | | | | | | XC = 1:

| | | | | | | | XD = 1: 0

| | | | | | | | XD=0: 1

| | | | | | | XC=0: 1

| | | | | | XG=0:

| | | | | | | XB = 1:

| | | | | | | | XD = 1:

| | | | | | | | | XE = 1: 1

| | | | | | | | | XE=0: 0

| | | | | | | | XD=0: 1

| | | | | | | XB=0: 0

| | | | | XS=0: 1

| | | | XJ=0:

| | | | | XC = 1:

| | | | | | XM = 1:

| | | | | | | XF = 1: 1

| | | | | | | XF=0:

| | | | | | | | XR = 1: 0

| | | | | | | | XR=0: 1

| | | | | | XM=0: 0

| | | | | XC=0:

| | | | | | XN = 1:

| | | | | | | XF = 1: 0

| | | | | | | XF=0:

| | | | | | | | XG = 1: 0

| | | | | | | | XG=0: 1

| | | | | | XN=0: 1

| | | XH=0:

| | | | XP = 1:

| | | | | XS = 1: 1

| | | | | XS=0:

| | | | | | XD = 1:

| | | | | | | XM = 1: 1

| | | | | | | XM=0:

| | | | | | | | XC = 1: 0

| | | | | | | | XC=0: 1

| | | | | | XD=0:

| | | | | | | XC = 1: 0

| | | | | | | XC=0:

| | | | | | | | XJ = 1: 1

| | | | | | | | XJ=0:

| | | | | | | | | XN = 1: 1

| | | | | | | | | XN=0: 0

| | | | XP=0:

| | | | | XF = 1:

| | | | | | XQ = 1:

| | | | | | | XK = 1: 1

| | | | | | | XK=0: 0

| | | | | | XQ=0:

| | | | | | | XK = 1:

| | | | | | | | XC = 1: 1

| | | | | | | | XC=0: 0

| | | | | | | XK=0: 1

| | | | | XF=0: 0

| XI=0:

| | XM = 1:

| | | XQ = 1: 0

| | | XQ=0:

| | | | XF = 1: 0

| | | | XF=0:

| | | | | XL = 1:

| | | | | | XC = 1: 1

| | | | | | XC=0:

| | | | | | | XB = 1:

| | | | | | | | XP = 1: 1

| | | | | | | | XP=0: 0

| | | | | | | XB=0: 0

| | | | | XL=0:

| | | | | | XC = 1:

| | | | | | | XH = 1:

| | | | | | | | XU = 1: 1

| | | | | | | | XU=0: 1

| | | | | | | XH=0: 1

| | | | | | XC=0: 1

| | XM=0:

| | | XQ = 1:

| | | | XJ = 1:

| | | | | XL = 1: 1

| | | | | XL=0:

| | | | | | XH = 1:

| | | | | | | XK = 1: 1

| | | | | | | XK=0:

| | | | | | | | XU = 1: 0

| | | | | | | | XU=0: 1

| | | | | | XH=0: 0

| | | | XJ=0:

| | | | | XN = 1: 0

| | | | | XN=0:

| | | | | | XP = 1:

| | | | | | | XB = 1: 0

| | | | | | | XB=0:

| | | | | | | | XF = 1: 1

| | | | | | | | XF=0: 0

| | | | | | XP=0: 1

| | | XQ=0:

| | | | XF = 1: 0

| | | | XF=0:

| | | | | XH = 1: 1

| | | | | XH=0:

| | | | | | XB = 1:

| | | | | | | XC = 1: 0

| | | | | | | XC=0: 1

| | | | | | XB=0: 0

XO=0:

| XM = 1:

| | XB = 1:

| | | XI = 1:

| | | | XC = 1: 0

| | | | XC=0:

| | | | | XK = 1: 0

| | | | | XK=0:

| | | | | | XP = 1:

| | | | | | | XS = 1: 0

| | | | | | | XS=0:

| | | | | | | | XG = 1:

| | | | | | | | | XF = 1: 1

| | | | | | | | | XF=0: 0

| | | | | | | | XG=0: 1

| | | | | | XP=0: 1

| | | XI=0: 0

| | XB=0:

| | | XF = 1:

| | | | XS = 1:

| | | | | XR = 1:

| | | | | | XI = 1:

| | | | | | | XG = 1: 1

| | | | | | | XG=0:

| | | | | | | | XQ = 1: 0

| | | | | | | | XQ=0:

| | | | | | | | | XJ = 1: 1

| | | | | | | | | XJ=0:

| | | | | | | | | | XD = 1: 1

| | | | | | | | | | XD=0: 0

| | | | | | XI=0:

| | | | | | | XU = 1: 0

| | | | | | | XU=0: 1

| | | | | XR=0:

| | | | | | XD = 1:

| | | | | | | XK = 1: 1

| | | | | | | XK=0:

| | | | | | | | XE = 1:

| | | | | | | | | XG = 1: 0

| | | | | | | | | XG=0: 1

| | | | | | | | XE=0: 0

| | | | | | XD=0: 1

| | | | XS=0:

| | | | | XQ = 1:

| | | | | | XI = 1:

| | | | | | | XL = 1:

| | | | | | | | XD = 1: 1

| | | | | | | | XD=0: 0

| | | | | | | XL=0: 1

| | | | | | XI=0: 1

| | | | | XQ=0:

| | | | | | XI = 1:

| | | | | | | XP = 1: 1

| | | | | | | XP=0: 0

| | | | | | XI=0: 0

| | | XF=0:

| | | | XG = 1:

| | | | | XD = 1:

| | | | | | XC = 1: 0

| | | | | | XC=0:

| | | | | | | XP = 1: 0

| | | | | | | XP=0:

| | | | | | | | XS = 1: 1

| | | | | | | | XS=0: 0

| | | | | XD=0: 1

| | | | XG=0: 0

| XM=0:

| | XF = 1: 0

| | XF=0:

| | | XB = 1:

| | | | XD = 1:

| | | | | XI = 1:

| | | | | | XG = 1: 0

| | | | | | XG=0: 1

| | | | | XI=0: 0

| | | | XD=0: 0

| | | XB=0:

| | | | XG = 1:

| | | | | XD = 1:

| | | | | | XP = 1: 0

| | | | | | XP=0:

| | | | | | | XE = 1: 1

| | | | | | | XE=0: 0

| | | | | XD=0:

| | | | | | XS = 1:

| | | | | | | XC = 1:

| | | | | | | | XH = 1: 1

| | | | | | | | XH=0: 0

| | | | | | | XC=0: 1

| | | | | | XS=0: 0

| | | | XG=0: 0

Test data from data set 1:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| No. | L | K | Infogain Accuracy | | Variance impurity | |
| Before Prune(%) | After Prune(%) | Before Prune(%) | After Prune(%) |
| 1 | 85 | 45 | 74.65 | 75.15 | 74.65 | 75.15 |
| 2 | 50 | 100 | 74.65 | 75.15 | 74.65 | 75.1 |
| 3 | 100 | 100 | 74.65 | 75.15 | 74.65 | 75.15 |
| 4 | 10 | 20 | 74.65 | 74.65 | 74.65 | 75.1 |
| 5 | 20 | 30 | 74.65 | 75.15 | 74.65 | 75.1 |
| 6 | 10 | 5 | 74.65 | 74.65 | 74.65 | 74.65 |
| 7 | 5 | 5 | 74.65 | 74.65 | 74.65 | 74.65 |
| 8 | 15 | 5 | 74.65 | 74.7 | 74.65 | 74.65 |
| 9 | 15 | 10 | 74.65 | 74.65 | 74.65 | 75.0 |
| 10 | 20 | 200 | 74.65 | 75.14 | 74.65 | 74.9 |

Test data from data set 2:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| No. | L | K | Infogain Accuracy | | Variance impurity | |
| Before Prune(%) | After Prune(%) | Before Prune(%) | After Prune(%) |
| 1 | 85 | 45 | 72.5 | 72.67 | 72.5 | 72.67 |
| 2 | 50 | 100 | 72.5 | 73.33 | 72.5 | 72.67 |
| 3 | 100 | 100 | 72.5 | 72.67 | 72.5 | 72.67 |
| 4 | 10 | 20 | 72.5 | 72.5 | 72.5 | 72.5 |
| 5 | 20 | 30 | 72.5 | 72.67 | 72.5 | 72.83 |
| 6 | 10 | 5 | 72.5 | 72.5 | 72.5 | 72.5 |
| 7 | 5 | 5 | 72.5 | 72.5 | 72.5 | 72.5 |
| 8 | 15 | 5 | 72.5 | 72.67 | 72.5 | 72.5 |
| 9 | 15 | 10 | 72.5 | 72.5 | 72.5 | 72.67 |
| 10 | 20 | 200 | 72.5 | 73.33 | 72.5 | 73.33 |