A)

1. Original Dataset Information Gain Tree (0.0 error rate)

```
***glucose
        +129
        ***tricepts
                +29
                        DECISION -> 1
                +28
                        DECISION -> 0
                +27
                        DECISION -> 0
                +26
                        DECISION -> 0
                +25
                        DECISION -> 0
                +24
                        DECISION -> 0
                +56
                        DECISION -> 0
                +23
                        DECISION -> 0
                +22
                        DECISION -> 0
                +54
                        DECISION -> 0
                +21
                        DECISION -> 0
                +20
                        DECISION -> 0
                +52
                        DECISION -> 0
                +51
                        DECISION -> 0
                +50
                        DECISION -> 0
                +8
                        DECISION -> 0
                +19
                        DECISION -> 0
                +7
                        DECISION -> 0
                +18
                        DECISION -> 0
                +17
                        DECISION -> 0
                +49
                        DECISION -> 1
                +16
                        DECISION -> 0
                +48
                        DECISION -> 0
                +15
                        DECISION -> 0
                +47
```

DECISION -> 0
+14
DECISION -> 0
+46
DECISION -> 1
+13
DECISION -> 0
+45
DECISION -> 0
+12
DECISION -> 0
+44
DECISION -> 0
+43
DECISION -> 0
+0
DECISION -> 0
+11
DECISION -> 0
+42
DECISION -> 0
+10
DECISION -> 0
+41
DECISION -> 0
+40
DECISION -> 0
+39
DECISION -> 0
+38
DECISION -> 0
+37
DECISION -> 0
+36
DECISION -> 1
+35
DECISION -> 0
+99
DECISION -> 0
+34
DECISION -> 0
+33
DECISION -> 0
+32
DECISION -> 0
+31
DECISION -> 0
+63
DECISION -> 0
+30
DECISION -> 1
+60
DECISION -> 0
+120
***pressure
+108
DECISION -> 0

+106
DECISION -> 0
+58
DECISION -> 0
+104
DECISION -> 0
+24
DECISION -> 0
+56
DECISION -> 0
+88
DECISION -> 0
+102
DECISION -> 0
+55
DECISION -> 0
+86
DECISION -> 1
+54
DECISION -> 0
+100
DECISION -> 0
+85
DECISION -> 0
+84
DECISION -> 0
+52
DECISION -> 0
+82
DECISION -> 0
+50
DECISION -> 0
+80
***pregnant
+9
DECISION -> 1
+8
DECISION -> 1
+17
DECISION -> 1
+7
DECISION -> 1
+6
DECISION -> 1
+15
DECISION -> 1
+5
DECISION -> 1
+14
DECISION -> 1
+4
DECISION -> 1
+13
DECISION -> 1
+3
DECISION -> 1
+12

DECISION -> 1
+2
DECISION -> 1
+11
DECISION -> 1
+10
DECISION -> 1
+1
DECISION -> 0
+0
DECISION -> 1

2. Original Dataset Gain Ratio Tree (0.0 error rate)

***glucose
+129
***pregnant
+9
DECISION -> 0
+8
DECISION -> 0
+17
DECISION -> 0
+7
DECISION -> 1
+6
DECISION -> 0
+15
DECISION -> 0
+5
DECISION -> 0
+14
DECISION -> 0
+4
DECISION -> 0
+13
DECISION -> 1
+3
DECISION -> 1
+12
DECISION -> 0
+2
DECISION -> 0
+11
DECISION -> 0
+10
***age
+29
DECISION -> 1
+28
DECISION -> 1
+27
DECISION -> 1
+59
DECISION -> 1
+26

DECISION -> 1
+58
DECISION -> 1
+25
DECISION -> 1
+57
DECISION -> 1
+24
DECISION -> 1
+56
DECISION -> 1
+23
DECISION -> 1
+55
DECISION -> 1
+22
DECISION -> 1
+54
DECISION -> 1
+21
DECISION -> 1
+53
DECISION -> 1
+52
DECISION -> 1
+51
DECISION -> 1
+50
DECISION -> 1
+81
DECISION -> 1
+49
DECISION -> 1
+48
DECISION -> 1
+47
DECISION -> 1
+46
DECISION -> 1
+45
DECISION -> 1
+44
DECISION -> 1
+43
DECISION -> 1
+42
DECISION -> 1
+41
DECISION -> 1
+40
DECISION -> 1
+72
DECISION -> 1
+70
DECISION -> 1
+39
DECISION -> 0

+38
DECISION -> 1
+37
DECISION -> 1
+69
DECISION -> 1
+36
DECISION -> 1
+68
DECISION -> 1
+35
DECISION -> 1
+67
DECISION -> 1
+34
DECISION -> 1
+66
DECISION -> 1
+33
DECISION -> 1
+65
DECISION -> 1
+32
DECISION -> 1
+64
DECISION -> 1
+31
DECISION -> 1
+63
DECISION -> 1
+30
DECISION -> 1
+62
DECISION -> 1
+61
DECISION -> 1
+60
DECISION -> 1
+126
***pregnant
+9
DECISION -> 0
+8
DECISION -> 0
+17
DECISION -> 0
+7
DECISION -> 0
+6
DECISION -> 0
+15
DECISION -> 0
+5
DECISION -> 0
+14
DECISION -> 0
+4

DECISION -> 0
+13
DECISION -> 1
+3
DECISION -> 0
+12
DECISION -> 0
+2
DECISION -> 0
+11
DECISION -> 0
+10
DECISION -> 0
+1
***age
+29
DECISION -> 1
+28
DECISION -> 1
+27
DECISION -> 1
+59
DECISION -> 1
+26
DECISION -> 1
+58
DECISION -> 1
+25
DECISION -> 1
+57
DECISION -> 1
+24
DECISION -> 1
+56
DECISION -> 1
+23
DECISION -> 1
+55
DECISION -> 1
+22
DECISION -> 1
+54
DECISION -> 1
+21
DECISION -> 0
+53
DECISION -> 1
+52
DECISION -> 1
+51
DECISION -> 1
+50
DECISION -> 1
+81
DECISION -> 1
+49
DECISION -> 1

+48
DECISION -> 1
+47
DECISION -> 1
+46
DECISION -> 1
+45
DECISION -> 1
+44
DECISION -> 1
+43
DECISION -> 1
+42
DECISION -> 1
+41
DECISION -> 1
+40
DECISION -> 1
+72
DECISION -> 1
+70
DECISION -> 1
+39
DECISION -> 1
+38
DECISION -> 1
+37
DECISION -> 1
+69
DECISION -> 1
+36
DECISION -> 1
+68
DECISION -> 1
+35
DECISION -> 1
+67
DECISION -> 1
+34
DECISION -> 1
+66
DECISION -> 1
+33
DECISION -> 1
+65
DECISION -> 1
+32
DECISION -> 1
+64
DECISION -> 1
+31
DECISION -> 1
+63
DECISION -> 1
+30
DECISION -> 1
+62

```
                    DECISION -> 1
            +61
                    DECISION -> 1
            +60
                    DECISION -> 1
    +0
            DECISION -> 0
```

2. Discretized Dataset Information Information Gain Tree (0.3291139240506329 error rate)

```
***glucose
    +129
    ***pregnant
        +9
            DECISION -> 1
        +8
            DECISION -> 1
        +7
            DECISION -> 1
        +6
            DECISION -> 1
        +5
            DECISION -> 1
        +14
            DECISION -> 1
        +4
            DECISION -> 1
        +13
            DECISION -> 1
        +12
            DECISION -> 1
        +3
            DECISION -> 1
        +11
            DECISION -> 1
        +2
            DECISION -> 1
        +10
            DECISION -> 0
        +1
            DECISION -> 1
        +0
            DECISION -> 1
    +134
    ***pregnant
        +9
            DECISION -> 0
        +8
            DECISION -> 1
        +7
            DECISION -> 1
        +6
            DECISION -> 1
        +5
            DECISION -> 1
        +14
```

DECISION -> 1
+4
DECISION -> 1
+13
DECISION -> 1
+12
DECISION -> 1
+3
DECISION -> 1
+11
DECISION -> 1
+2
DECISION -> 1
+10
DECISION -> 1
+1
DECISION -> 1
+0
DECISION -> 1

4. Discretized Dataset Gain Ratio Tree (0.275 error rate)

***pregnant
+9
***glucose
+129
DECISION -> 1
+128
DECISION -> 1
+127
DECISION -> 1
+126
DECISION -> 1
+125
DECISION -> 1
+189
DECISION -> 1
+124
DECISION -> 1
+123
DECISION -> 0
+187
DECISION -> 1
+122
DECISION -> 1
+121
DECISION -> 1
+120
DECISION -> 1
+184
DECISION -> 1
+183
DECISION -> 1
+180
DECISION -> 1
+119

DECISION -> 1
+118
DECISION -> 1
+117
DECISION -> 1
+115
DECISION -> 1
+114
DECISION -> 1
+112
DECISION -> 1
+111
DECISION -> 1
+176
DECISION -> 1
+175
DECISION -> 1
+174
DECISION -> 1
+173
DECISION -> 1
+171
DECISION -> 1
+109
DECISION -> 1
+108
DECISION -> 1
+107
DECISION -> 1
+106
DECISION -> 1
+105
DECISION -> 1
+104
DECISION -> 1
+169
DECISION -> 1
+103
DECISION -> 1
+168
DECISION -> 1
+102
DECISION -> 1
+101
DECISION -> 1
+166
DECISION -> 1
+100
DECISION -> 1
+164
DECISION -> 1
+163
DECISION -> 1
+162
DECISION -> 1
+161
DECISION -> 1

+99
	DECISION -> 1
+98
	DECISION -> 1
+96
	DECISION -> 1
+95
	DECISION -> 1
+94
	DECISION -> 1
+93
	DECISION -> 1
+92
	DECISION -> 1
+91
	DECISION -> 1
+90
	DECISION -> 1
+159
	DECISION -> 1
+158
	DECISION -> 1
+156
	DECISION -> 1
+154
	DECISION -> 1
+153
	DECISION -> 1
+152
	DECISION -> 1
+151
	DECISION -> 1
+150
	DECISION -> 1
+89
	DECISION -> 1
+88
	DECISION -> 1
+87
	DECISION -> 1
+86
	DECISION -> 1
+84
	DECISION -> 1
+83
	DECISION -> 1
+82
	DECISION -> 1
+80
	DECISION -> 1
+148
	DECISION -> 1
+147
	DECISION -> 1
+146
	DECISION -> 1
+145

DECISION -> 1
+143
DECISION -> 1
+141
DECISION -> 1
+140
DECISION -> 1
+79
DECISION -> 1
+78
DECISION -> 1
+75
DECISION -> 1
+74
DECISION -> 1
+73
DECISION -> 1
+71
DECISION -> 1
+139
DECISION -> 1
+138
DECISION -> 1
+137
DECISION -> 1
+136
DECISION -> 1
+135
DECISION -> 1
+134
DECISION -> 0
+198
DECISION -> 1
+197
DECISION -> 1
+196
DECISION -> 1
+130
DECISION -> 1
+67
DECISION -> 1
+190
DECISION -> 1
+62
DECISION -> 1
+8
***age
+29
DECISION -> 1
+28
DECISION -> 1
+27
DECISION -> 1
+26
DECISION -> 1
+58
DECISION -> 1

+25
DECISION -> 1
+57
DECISION -> 1
+24
DECISION -> 1
+23
DECISION -> 1
+55
DECISION -> 1
+22
DECISION -> 1
+54
DECISION -> 1
+21
DECISION -> 1
+52
DECISION -> 1
+51
DECISION -> 1
+50
DECISION -> 1
+81
DECISION -> 1
+48
DECISION -> 1
+47
DECISION -> 1
+46
DECISION -> 1
+45
DECISION -> 1
+44
DECISION -> 1
+43
DECISION -> 1
+42
DECISION -> 1
+41
DECISION -> 1
+40
DECISION -> 1
+39
DECISION -> 0
+38
DECISION -> 1
+37
DECISION -> 1
+36
DECISION -> 1
+35
DECISION -> 1
+67
DECISION -> 1
+34
DECISION -> 0
+66

DECISION -> 1
+65
DECISION -> 1
+33
DECISION -> 1
+32
DECISION -> 1
+63
DECISION -> 1
+31
DECISION -> 1
+30
DECISION -> 1
+62
DECISION -> 1

B)

Original DataSet

Info gain: 0.0

Gain ratio: 0.0

Discretized DataSet

Info gain: 0.3291139240506329

Gain ratio: 0.275

C)

(i) Does DTL with Information Gain behave better or worse than DTL with Gain Ratio? Consider performances to be the same if they are within 1% of each other (this is a shortcut to statistical significance).

Answer:

Information gain did better when you look at error rate but creates a significantly larger tree and take much more computational power to test than w/ Gain ratio.

(ii) A priori, what algorithm would you have expected to perform the best and why? Did the results confirm your expectations?

Answer:

Gain ratio seems to perform the best since it usually makes for a smaller tree and than information gain and also creates a lower error rate.