Problem 1

The tree before pruning:

node 0

Top: 6,4, 0.97

Education Level gain = 0.125

Career gain = 0.125

Years of Experience gain = 0.020

Selected Attribute: Education

node 1

High School: 4,1, 0.72

Career gain = 0.170

Years of Experience gain = 0.322

Selected Attribute: Years of Experience

node 3

More than 10: 1, 1, 1.00

Career gain = 1.00

Selected Attribute: Career

node 8

Management, 1, 0

Class: High

node 9

Service: 1,0

Class Low

node 4

Less than 3: 1, 0 ,0.00

Career gain = 0.00

Selected Attribute: Career

node 10

Management: 1, 0

Class Low

node 5

3 to 10: 1, 0

Career gain = 0.00

Selected Attribute: Career

node 11

Management: 1, 0

Class Low

node 12

Service: 1, 0

Class Low

node 2

College 3,2, 0.97

Career gain = 0.420

Years of Experience gain = 0.171

Selected Attribute: Career

node 6

Management 2,0

Years of Experience gain = 0.00

node 13

More than 10

Class High

node 14

Less than 3

Class High

node 7

Service 2,1, 0.91

node 15

More than 10

Class Low

node 16

Less than 3

Class Low

node 17

3 to 10

Class High

for the 1 iteration:

node num: sum\_of\_total\_wrong: wrong\_count:

node 17 0 1

node 16 0 0

node 15 0 0

node 14 0 1

node 13 0 0

node 12 0 0

node 11 0 0

node 10 0 0

node 9 0 0

node 8 0 0

node 7 1 0

node 7 need to be pruned

node 7 will be labeled with Low class

for the 2 iteration:

node num: sum\_of\_total\_wrong: wrong\_count:

node 14 0 1

node 13 0 0

node 12 0 0

node 11 0 0

node 10 0 0

node 9 0 0

node 8 0 0

node 7 0 0

node 6 1 1

node 5 0 0

node 4 0 0

node 3 0 0

node 2 1 2

node 1 0 1

node 0 1 1

The final tree:

node 0

Top: 6,4, 0.97

Education Level gain = 0.125

Career gain = 0.125

Years of Experience gain = 0.020

Selected Attribute: Education

node 1

High School: 4,1, 0.72

Career gain = 0.170

Years of Experience gain = 0.322

Selected Attribute: Years of Experience

node 3

More than 10: 1, 1, 1.00

Career gain = 1.00

Selected Attribute: Career

node 8

Management, 1, 0

Class: High

node 9

Service: 1,0

Class Low

node 4

Less than 3: 1, 0 ,0.00

Career gain = 0.00

Selected Attribute: Career

node 10

Management: 1, 0

Class Low

node 5

3 to 10: 1, 0

Career gain = 0.00

Selected Attribute: Career

node 11

Management: 1, 0

Class Low

node 12

Service: 1, 0

Class Low

node 2

College 3,2, 0.97

Career gain = 0.420

Years of Experience gain = 0.171

Selected Attribute: Career

node 6

Management 2,0

Years of Experience gain = 0.00

node 13

More than 10

Class High

node 14

Less than 3

Class High

node 7

Service 2,1, 0.91

Class Low

Problem2

6.7 3.1 4.4 1.4 x1 versicolor versicolor versicolor versicolor versicolor

4.4 3.2 1.3 0.2 x2 virginica setosa setosa setosa setosa

5.3 3.7 1.5 0.2 x3 setosa setosa setosa setosa setosa

7.7 2.8 6.7 2.0 x10 virginica virginica virginica virginica virginica

5.1 3.5 1.4 0.2 x11 setosa setosa setosa setosa setosa

6.5 3.0 5.2 2.0 x12 virginica virginica virginica virginica virginica

7.1 3.0 5.9 2.1 x13 virginica virginica virginica virginica virginica

6.4 2.7 5.3 1.9 x14 virginica virginica virginica virginica virginica

5.2 2.7 3.9 1.4 x15 versicolor versicolor versicolor versicolor versicolor

7.0 3.2 4.7 1.4 x16 versicolor versicolor versicolor versicolor versicolor

7.2 3.2 6.0 1.8 x17 virginica virginica virginica virginica virginica

5.4 3.7 1.5 0.2 x19 setosa setosa setosa setosa setosa

5.6 3.0 4.5 1.5 x20 versicolor versicolor versicolor versicolor versicolor

5.9 3.2 4.8 1.8 x21 versicolor versicolor versicolor versicolor versicolor

5.1 3.4 1.5 0.2 x22 setosa setosa setosa setosa setosa

6.9 3.1 4.9 1.5 x23 versicolor versicolor virginica versicolor virginica

6.0 2.2 4.0 1.0 x24 versicolor versicolor versicolor versicolor versicolor

4.7 3.2 1.6 0.2 x25 virginica setosa setosa setosa setosa

4.6 3.6 1.0 0.2 x27 setosa setosa setosa setosa setosa

5.6 3.0 4.1 1.3 x31 versicolor versicolor versicolor versicolor versicolor

5.5 3.5 1.3 0.2 x33 setosa setosa setosa setosa setosa

5.5 2.4 3.8 1.1 x34 versicolor versicolor versicolor versicolor versicolor

5.1 3.8 1.6 0.2 x35 setosa setosa setosa setosa setosa

6.3 3.3 4.7 1.6 x36 virginica versicolor versicolor versicolor versicolor

6.6 2.9 4.6 1.3 x100 versicolor versicolor versicolor versicolor versicolor

7.7 3.0 6.1 2.3 x101 virginica virginica virginica virginica virginica

6.4 2.9 4.3 1.3 x102 versicolor versicolor versicolor versicolor versicolor

6.9 3.1 5.1 2.3 x103 virginica virginica virginica virginica virginica

6.7 3.0 5.0 1.7 x104 virginica versicolor versicolor virginica virginica

4.3 3.0 1.1 0.1 x105 setosa setosa setosa setosa setosa

7.7 2.6 6.9 2.3 x106 virginica virginica virginica virginica virginica

6.7 3.3 5.7 2.5 x107 virginica virginica virginica virginica virginica

6.7 2.5 5.8 1.8 x108 virginica virginica virginica virginica virginica

1.0 3.1 1.6 0.2 x109 setosa setosa setosa setosa setosa

5.7 4.4 1.5 0.4 x110 setosa setosa setosa setosa setosa

6.5 3.0 5.5 1.8 x111 virginica virginica virginica virginica virginica

6.1 3.0 4.9 1.8 x112 versicolor virginica virginica virginica versicolor

5.4 3.4 1.7 0.2 x113 setosa setosa setosa setosa setosa

6.5 3.2 5.1 2.0 x114 virginica virginica virginica virginica virginica

5.2 3.4 1.4 0.2 x115 setosa setosa setosa setosa setosa

5.7 3.0 4.2 1.2 x116 versicolor versicolor versicolor versicolor versicolor

5.5 2.3 4.0 1.3 x117 versicolor versicolor versicolor versicolor versicolor

5.0 3.4 1.6 0.4 x118 setosa setosa setosa setosa setosa

5.8 2.7 5.1 1.9 x119 virginica virginica versicolor virginica virginica

6.1 2.8 4.0 1.3 x120 versicolor versicolor versicolor versicolor versicolor

5.7 2.5 5.0 2.0 x121 virginica virginica virginica virginica virginica

6.3 2.9 5.6 1.8 x122 virginica virginica virginica virginica virginica

4.9 3.1 1.5 0.1 x123 setosa setosa setosa setosa setosa

6.8 3.2 5.9 2.3 x124 virginica virginica virginica virginica virginica

6.9 3.2 5.7 2.3 x125 virginica virginica virginica virginica virginica

6.7 3.1 4.7 1.5 x126 versicolor versicolor versicolor versicolor versicolor

5.7 2.8 4.1 1.3 x127 versicolor versicolor versicolor versicolor versicolor

5.0 3.5 1.6 0.6 x128 setosa setosa setosa setosa setosa

5.4 3.9 1.7 0.4 x129 setosa setosa setosa setosa setosa

5.2 3.5 1.5 0.2 x130 setosa setosa setosa setosa setosa

6.1 2.8 4.7 1.2 x131 versicolor versicolor versicolor versicolor versicolor

5.7 2.9 4.2 1.3 x132 versicolor versicolor versicolor versicolor versicolor

5.8 2.7 3.9 1.2 x133 versicolor versicolor versicolor versicolor versicolor

5.0 3.3 1.4 0.2 x134 setosa setosa setosa setosa setosa

6.8 2.8 4.8 1.4 x135 versicolor versicolor versicolor versicolor versicolor

6.3 2.8 5.1 1.5 x136 versicolor versicolor versicolor versicolor versicolor

6.2 2.2 4.5 1.5 x137 versicolor versicolor versicolor versicolor versicolor

6.0 3.0 4.8 1.8 x138 versicolor versicolor virginica versicolor versicolor

5.1 3.5 1.4 0.3 x139 setosa setosa setosa setosa setosa

5.7 3.8 1.7 0.3 x140 setosa setosa setosa setosa setosa

6.1 3.0 4.6 1.4 x141 versicolor versicolor versicolor versicolor versicolor

5.8 4.0 1.2 0.2 x142 setosa setosa setosa setosa setosa

7.2 3.6 6.1 2.5 x143 virginica virginica virginica virginica virginica

6.1 2.6 5.6 1.4 x144 versicolor versicolor versicolor versicolor virginica

5.5 2.5 4.0 1.3 x145 versicolor versicolor versicolor versicolor versicolor

7.3 2.9 6.3 1.8 x146 virginica virginica virginica virginica virginica

4.8 3.0 1.4 0.1 x147 setosa setosa setosa setosa setosa

7.6 3.0 6.6 2.1 x148 virginica virginica virginica virginica virginica

6.5 3.0 5.8 2.2 x149 virginica virginica virginica virginica virginica

5.1 3.3 1.7 0.5 x150 setosa setosa setosa setosa setosa

Problem 3

==========

poly-kernel

exponent = 1

result:

=== Confusion Matrix ===

a b <-- classified as

356 73 | a = car

56 361 | b = noncar

correct 717

incorect 129

==========

poly-kernel

exponent = 2

result:

=== Confusion Matrix ===

a b <-- classified as

408 21 | a = car

15 402 | b = noncar

correct 810

incorrect 36

==========

poly-kernel

exponent = 4

result:

=== Confusion Matrix ===

a b <-- classified as

401 28 | a = car

27 390 | b = noncar

correct 791

incorrect 55

==========

rbf kernel

gamma = 0.01

result:

=== Confusion Matrix ===

a b <-- classified as

267 162 | a = car

70 347 | b = noncar

correct 614

incorrect 232

==========

rbf kernel

gamma = 1

result:

=== Confusion Matrix ===

a b <-- classified as

373 56 | a = car

26 391 | b = noncar

correct 764

incorrect 82

From the result, poly kernel, when exponent equals 2, the number of correct instances is greater than the number of correct instances when exponent equals to 1 and 4. For rbf kernel when gamma equals to 1 the number of correct instances is greater than the number of correct instances when gamma equals to 0.01

The kernel function map the original data set to a high dimension data set. If the higher dimension data set could be separated, the SVM would achieve a better result. So for the poly kernel, when exponent equals to 2 the mapped data is more separable than the mapped data when exponent equals to 1 and 4. For the same reason, when gamma eqs to 1 the mapped data is more separable than the mapped data when gamma eqs to 0.01

Problem 4

Let support Φ(x) = Φ(x1, x2) = x1 + ex2 , then K (x, z) = K (x1, x2, z1, z2) =Φ(x) \*Φ(z). From the definition of kernel we can see that K (x, x) =Φ(x)2 . So that K (x, z) = x1 \* z1 + x1 \* ez2 + z1 \* ex2 + ex2 + z2   is kernel.

Problem 5

Because a0 = -0.8, a1 = 1, a2 = 6.4, a3 = -1.9, so [1,1,1] and [1,1,0] are support vectors.

Because Macintosh HD:Users:bohan:Desktop:Screen Shot 2015-02-18 at 12.03.00 PM.png, so that w = [5,4, -1, -1]T , and according to Macintosh HD:Users:bohan:Desktop:Screen Shot 2015-02-18 at 12.04.09 PM.png, so b = -4.4.

To classify [1, 0.8, 1], we apply the formula wT \* x + b, which equals to -0.8. So we can conclude that even the tuple [1, 0.8, 1] lies between the support vectors, it still closer to the negative class. [1, 0.8, 1] is labeled with class -1.