ASSIGNMENT 7

Probabilistic Approaches

Chunyao Wang

Wenxing Li

Jie Rong

1 Classification of iris flowers

In this part, we implement different classification method and parameters to classify the flower iris in the file “iris.arff” .

1.1 BFTree

With default parameters, we get the accuracy 94.7%.

```
Correctly Classified Instances         142               94.6667 %
Incorrectly Classified Instances         8                5.3333 %
```

We try different parameters to obtain better results. We find when we set the minNumObj as 3, the results become better. The accuracy is 96%.

```
Correctly Classified Instances         144               96      %
Incorrectly Classified Instances         6                4      %
```

The classification rules are shown below:

```
Best-First Decision Tree

PL < 2.45: Iris_setosa(50.0/0.0)
PL >= 2.45
|  PW < 1.75
|  |  PL < 4.95: Iris_versicolor(47.0/1.0)
|  |  PL >= 4.95: Iris_virginica(4.0/2.0)
|  PW >= 1.75: Iris_virginica(45.0/1.0)
```

1.2 FT

With default parameters, we get the accuracy 96.7%. When we use difference parameters, the result does not become better.

```
Correctly Classified Instances         145               96.6667 %
Incorrectly Classified Instances         5                3.3333 %
```

The classification rules is shown below:

```
=== Classifier model (full training set) ===

FT tree
------------------
: FT_1:15/15 (150)

Number of Leaves  :     1

Size of the Tree :      1
FT_1:
Class 0 :
23.67 +
[PL] * -7.42 +
[PW] * -5.71

Class 1 :
-5.82 +
[SL] * 1.67 +
[SW] * 0.11 +
[PL] * -0.4 +
[PW] * -1.28

Class 2 :
-31.02 +
[SL] * -0.4 +
[SW] * -3.32 +
[PL] * 5.56 +
[PW] * 9.9
```

1.3 J48

With default parameters, we get the accuracy 96%. When we use difference parameters, the result does not become better.

```
Correctly Classified Instances         144               96      %
Incorrectly Classified Instances         6                4      %
```

The classification rules are shown below:

```
J48 pruned tree
------------------

PW <= 0.6: Iris_setosa (50.0)
PW > 0.6
|   PW <= 1.7
|   |   PL <= 4.9: Iris_versicolor (48.0/1.0)
|   |   PL > 4.9
|   |   |   PW <= 1.5: Iris_virginica (3.0)
|   |   |   PW > 1.5: Iris_versicolor (3.0/1.0)
|   PW > 1.7: Iris_virginica (46.0/1.0)
```

1.4 LADTree

With default parameters, we get the accuracy 94%.

```
Correctly Classified Instances         141               94      %
Incorrectly Classified Instances         9                6      %
```

The accuracy can be improved to 95.3% if we change numOfBoostingIterations from 10 to 30.

```
Correctly Classified Instances         143               95.3333 %
Incorrectly Classified Instances         7                4.6667 %
```

The classification rules are shown below:

```
weka.classifiers.trees.LADTree:

: 0,0,0
|  (1)PW < 0.8: 2,-1,-1
|  (1)PW >= 0.8: -1,0.5,0.5
|  |  (2)PW < 1.75: -0.584,1.39,-0.805
|  |  |  (16)PL < 5.35: -0.445,0.669,-0.224
|  |  |  (16)PL >= 5.35: -0.444,-0.448,0.892
|  |  (2)PW >= 1.75: -0.584,-0.996,1.58
|  |  (10)PL < 4.45: -0.443,1.055,-0.612
|  |  (10)PL >= 4.45: -0.457,-0.185,0.642
|  |  (13)PL < 4.85: -0.448,0.687,-0.239
|  |  (13)PL >= 4.85: -0.448,-0.089,0.537
|  |  (14)PL < 4.95: -0.89,1.287,-0.397
|  |  (14)PL >= 4.95: -0.89,-0.267,1.158
|  |  |  (15)PW < 1.55: -0.444,-0.449,0.893
|  |  |  (15)PW >= 1.55: -0.445,0.567,-0.122
|  |  |  |  (17)PW < 1.75: -0.444,0.774,-0.329
|  |  |  |  |  (18)SL < 6.95: -0.444,0.89,-0.445
|  |  |  |  |  (18)SL >= 6.95: -0.444,-0.446,0.89
|  |  |  |  (17)PW >= 1.75: -0.444,-0.445,0.889
|  (3)PL < 4.95: 0.355,0.256,-0.611
|  |  (4)PL < 2.45: 0.931,-0.474,-0.457
|  |  (4)PL >= 2.45: -0.539,0.572,-0.033
|  |  |  (5)SL < 4.95: -0.488,-2.064,2.552
|  |  |  (5)SL >= 4.95: -0.501,0.471,0.03
|  |  |  (6)PW < 1.65: -1.322,3.015,-1.694
|  |  |  (6)PW >= 1.65: -1.352,-0.911,2.263
|  |  |  |  (7)SW < 3.1: -1.782,-1.896,3.678
|  |  |  |  (7)SW >= 3.1: -1.78,5.609,-3.828
|  (3)PL >= 4.95: -0.447,-0.858,1.305
|  |  (8)PW < 1.55: -0.476,-1.117,1.593
|  |  (8)PW >= 1.55: -0.497,0.279,0.218
|  |  |  (9)PL < 5.15: -0.917,1.91,-0.992
|  |  |  |  (11)PW < 1.75: -1.335,2.925,-1.589
|  |  |  |  (11)PW >= 1.75: -1.334,-1.358,2.692
|  |  |  (9)PL >= 5.15: -0.883,-1.232,2.115
```

```
|   (12)PL < 2.45: 2.681,-1.343,-1.338
|   (12)PL >= 2.45: -1.352,0.406,0.946
|   (19)PW < 1.35: 0.875,-0.417,-0.458
|   (19)PW >= 1.35: -0.445,-0.076,0.52
```

1.5 LMT

With default parameters, we get the accuracy 94%.

```
Correctly Classified Instances         141               94      %
Incorrectly Classified Instances         9                6      %
```

Then We change numOfBoostingIterations from -1 to 15 and minNumInstances from 15 to 39, the result becomes better. The accuracy is 98%.

```
Correctly Classified Instances         147               98      %
Incorrectly Classified Instances         3                2      %
```

The classification rules are shown below:

```
Logistic model tree
------------------
: LM_1:15/15 (150)

Number of Leaves  :     1

Size of the Tree :      1
LM_1:
Class 0 :
23.67 +
[PL] * -7.42 +
[PW] * -5.71

Class 1 :
-5.82 +
[SL] * 1.67 +
[SW] * 0.11 +
[PL] * -0.4 +
[PW] * -1.28

Class 2 :
-31.02 +
[SL] * -0.4 +
[SW] * -3.32 +
[PL] * 5.56 +
[PW] * 9.9
```

1.6 NBTree

With default parameters, we get the accuracy 94.7%. This method does not have other parameters.

```
Correctly Classified Instances         142               94.6667 %
Incorrectly Classified Instances         8                5.3333 %
```

The classification rules are shown below:

```
NBTree
------------------

SW <= 3.35
|   SW <= 2.95
|   |   PL <= 4.75: NB 3
|   |   PL > 4.75: NB 4
|   SW > 2.95: NB 5
SW > 3.35
|   SL <= 5.9: NB 7
|   SL > 5.9: NB 8
```

1.7 RandomTree

We get the accuracy 91.3% using default parameters.

```
Correctly Classified Instances         137               91.3333 %
Incorrectly Classified Instances        13                8.6667 %
```

Then we change maxDepth from 0 to 10, minNum from 1.0 to 8.0 and seed from 1 to 10, the accuracy is improved to 95.3%.

```
Correctly Classified Instances         143               95.3333 %
Incorrectly Classified Instances         7                4.6667 %
```

The classification rules are shown below:

```
RandomTree
==========

PL < 2.45 : Iris_setosa (50/0)
PL >= 2.45
|   PW < 1.75
|   |   PL < 4.95
|   |   |   PW < 1.65 : Iris_versicolor (47/0)
|   |   |   PW >= 1.65 : Iris_virginica (1/0)
|   |   PL >= 4.95 : Iris_virginica (6/2)
|   PW >= 1.75
|   |   PL < 4.85 : Iris_virginica (3/1)
|   |   PL >= 4.85 : Iris_virginica (43/0)
```

1.8 REPTree

We get the accuracy 94% using default parameters.

```
Correctly Classified Instances         141               94      %
Incorrectly Classified Instances         9                6      %
```

When we try difference parameters, the accuracy does not increase.

The classification rules are shown below:

```
REPTree
============

PL < 2.5 : Iris_setosa (33/0) [17/0]
PL >= 2.5
|   PW < 1.75 : Iris_versicolor (36/3) [18/2]
|   PW >= 1.75 : Iris_virginica (31/1) [15/0]
```

1.9 SimpleCart

We get the accuracy of 95.3% using default parameters.

```
Correctly Classified Instances         143               95.3333 %
Incorrectly Classified Instances         7                4.6667 %
```

Other different parameters cannot make accuracy better.

The classification rules are shown below:

```
CART Decision Tree

PL < 2.45: Iris_setosa(50.0/0.0)
PL >= 2.45
|  PW < 1.75
|  |  PL < 4.95: Iris_versicolor(47.0/1.0)
|  |  PL >= 4.95
|  |  |  PW < 1.55: Iris_virginica(3.0/0.0)
|  |  |  PW >= 1.55: Iris_versicolor(2.0/1.0)
|  PW >= 1.75: Iris_virginica(45.0/1.0)
```

1.10 summary

According to all analysis above, it turns out that we get the best accuracy when we implement the LMT method. Hence the highest accuracy we can get is 98%.

2 Classification of Congressmen

In this part, we implement different classification method and parameters to classify the congressmen in the file “house-votes-84.arff” .

2.1 ADTree

Using default parameters, we can get the accuracy 95.86%.

```
Correctly Classified Instances         417               95.8621 %
Incorrectly Classified Instances        18                4.1379 %
```

After we change numOfBoostingIterations from 10 to 25, the result becomes better. The accuracy is improved to 96.32%.

```
Correctly Classified Instances         419               96.3218 %
Incorrectly Classified Instances        16                3.6782 %
```

The classification rules are shown below:

```
Alternating decision tree:

: 0.231
|  (1)physician_fee_freeze = y: -1.417
|  |  (5)adoption_of_the_budget_resolution = n: -0.558
|  |  |  (8)superfund_right_to_sue = y: -0.62
|  |  |  |  (16)water_project_cost_sharing = n: 0.14
|  |  |  |  (16)water_project_cost_sharing != n: -0.688
|  |  |  |  |  (17)export_administration_act_sa = w: 0.281
|  |  |  |  |  (17)export_administration_act_sa != w: -0.619
|  |  |  (8)superfund_right_to_sue != y: 0.59
|  |  (5)adoption_of_the_budget_resolution != n: 0.518
|  |  |  (7)nti_satellite_test_ban = n: 0.984
|  |  |  |  (12)water_project_cost_sharing = y: 0.79
|  |  |  |  (12)water_project_cost_sharing != y: -0.469
|  |  |  (7)nti_satellite_test_ban != n: -0.981
|  |  |  (19)superfund_right_to_sue = y: 0.359
|  |  |  (19)superfund_right_to_sue != y: -0.518
|  |  (6)immigration = n: 0.442
|  |  (6)immigration != n: -0.829
|  |  (13)synfuels_corporation_cutback = y: 0.156
|  |  (13)synfuels_corporation_cutback != y: -0.52
|  |  (20)export_administration_act_sa = y: -0.416
|  |  (20)export_administration_act_sa != y: 0.19
|  (1)physician_fee_freeze != y: 1.66
|  |  (2)adoption_of_the_budget_resolution = y: 2.381
|  |  (2)adoption_of_the_budget_resolution != y: -0.956
|  (3)synfuels_corporation_cutback = y: 0.914
|  |  (9)export_administration_act_sa = y: 0.011
|  |  (9)export_administration_act_sa != y: 0.791
|  (3)synfuels_corporation_cutback != y: -0.566
|  |  (21)water_project_cost_sharing = n: -0.363
|  |  (21)water_project_cost_sharing != n: 0.161
|  (4)education_spending = n: 0.684
|  |  (11)duty_free_exports = n: -0.314
|  |  (11)duty_free_exports != n: 0.558
|  |  (15)immigration = y: -0.139
```

```
|   |   (15)immigration != y: 0.514
|   |   |   (18)aid_to_nicaraguan_contras = w: -0.229
|   |   |   (18)aid_to_nicaraguan_contras != w: 0.526
|   (4)education_spending != n: -0.346
|   |   (10)physician_fee_freeze = n: 0.614
|   |   |   (14)education_spending = y: 0.682
|   |   |   (14)education_spending != y: -0.392
|   |   (10)physician_fee_freeze != n: -0.381
|   (22)mx_missile = y: 0.344
|   |   (24)duty_free_exports = n: -0.249
|   |   (24)duty_free_exports != n: 0.554
|   (22)mx_missile != y: -0.119
|   |   (23)education_spending = w: -0.619
|   |   (23)education_spending != w: 0.187
```

2.2 BFTree

With default parameters, we get the accuracy of 95.17%.

```
Correctly Classified Instances         414               95.1724 %
Incorrectly Classified Instances        21                4.8276 %
```

The accuracy increases to 95.86% when we change the minNumObj from 2 to 5 and seed from 1 to 2.

```
Correctly Classified Instances         417               95.8621 %
Incorrectly Classified Instances        18                4.1379 %
```

The classification rules are shown below:

```
Best-First Decision Tree

physician_fee_freeze=(n)|(w): democrat(253.0/5.0)
physician_fee_freeze!=(n)|(w): republican(163.0/14.0)
```

2.3 DecisionStump

We get the accuracy 95.63 using default parameters. This method does not have other parameters.

```
Correctly Classified Instances         416               95.6322 %
Incorrectly Classified Instances        19                4.3678 %
```

The classification rules are shown below:

```
Decision Stump

Classifications

physician_fee_freeze = y : republican
physician_fee_freeze != y : democrat
physician_fee_freeze is missing : democrat

Class distributions

physician_fee_freeze = y
republican      democrat
0.9209039548022598      0.07909604519774012
physician_fee_freeze != y
republican      democrat
0.01937984496124031     0.9806201550387597
physician_fee_freeze is missing
republican      democrat
0.38620689655172413     0.6137931034482759
```

2.4 FT

We get the accuracy 96.78% with default parameters.

```
Correctly Classified Instances         421               96.7816 %
Incorrectly Classified Instances        14                3.2184 %
```

The classification rules are shown below:

```
FT tree
------------------

N0#1 <= 0.595535: FT_1:15/30 (268)
N0#1 > 0.595535: FT_2:15/30 (167)
```

2.5 Id3

The accuracy is 94.25% with default parameters.

```
Correctly Classified Instances         410               94.2529 %
Incorrectly Classified Instances        22                5.0575 %
```

The classification rules are shown below:

```
Id3


physician_fee_freeze = y
|  synfuels_corporation_cutback = w: republican
|  synfuels_corporation_cutback = n
|  |  duty_free_exports = n
|  |  |  adoption_of_the_budget_resolution = n: republican
|  |  |  adoption_of_the_budget_resolution = y
|  |  |  |  export_administration_act_sa = y: republican
|  |  |  |  export_administration_act_sa = w
|  |  |  |  |  handicapped_infants = n: democrat
|  |  |  |  |  handicapped_infants = w: null
|  |  |  |  |  handicapped_infants = y: republican
|  |  |  |  export_administration_act_sa = n: null
|  |  |  adoption_of_the_budget_resolution = w: null
|  |  duty_free_exports = y
|  |  |  immigration = y: republican
|  |  |  immigration = n
|  |  |  |  export_administration_act_sa = y: democrat
|  |  |  |  export_administration_act_sa = w
|  |  |  |  |  water_project_cost_sharing = y: republican
|  |  |  |  |  water_project_cost_sharing = n: democrat
|  |  |  |  |  water_project_cost_sharing = w: null
|  |  |  |  export_administration_act_sa = n: republican
|  |  |  immigration = w: null
|  |  duty_free_exports = w: republican
|  synfuels_corporation_cutback = y
|  |  adoption_of_the_budget_resolution = n
|  |  |  el_salvador_aid = y
|  |  |  |  export_administration_act_sa = y: republican
|  |  |  |  export_administration_act_sa = w
|  |  |  |  |  handicapped_infants = n: republican
|  |  |  |  |  handicapped_infants = w: null
|  |  |  |  |  handicapped_infants = y: democrat
|  |  |  |  export_administration_act_sa = n
|  |  |  |  |  superfund_right_to_sue = y
|  |  |  |  |  |  water_project_cost_sharing = y: republican
|  |  |  |  |  |  water_project_cost_sharing = n
|  |  |  |  |  |  |  handicapped_infants = n: democrat
|  |  |  |  |  |  |  handicapped_infants = w: null
|  |  |  |  |  |  |  handicapped_infants = y: republican
|  |  |  |  |  |  water_project_cost_sharing = w: null
|  |  |  |  |  superfund_right_to_sue = w: null
|  |  |  |  |  superfund_right_to_sue = n: democrat
|  |  |  el_salvador_aid = w: null
|  |  |  el_salvador_aid = n: democrat
|  |  adoption_of_the_budget_resolution = y
|  |  |  nti_satellite_test_ban = n: democrat
|  |  |  nti_satellite_test_ban = y: republican
|  |  |  nti_satellite_test_ban = w: null
|  |  adoption_of_the_budget_resolution = w: democrat
physician_fee_freeze = w
|  mx_missile = n: democrat
|  mx_missile = y
|  |  nti_satellite_test_ban = n: republican
|  |  nti_satellite_test_ban = y: democrat
|  |  nti_satellite_test_ban = w: democrat
|  mx_missile = w: republican
```

```
physician_fee_freeze = n
|  adoption_of_the_budget_resolution = n
|  |  education_spending = y: democrat
|  |  education_spending = n
|  |  |  synfuels_corporation_cutback = w: null
|  |  |  synfuels_corporation_cutback = n
|  |  |  |  religious_groups_in_schools = y: democrat
|  |  |  |  religious_groups_in_schools = n
|  |  |  |  |  crime = y: republican
|  |  |  |  |  crime = n: democrat
|  |  |  |  |  crime = w: null
|  |  |  |  religious_groups_in_schools = w: null
|  |  |  synfuels_corporation_cutback = y: democrat
|  |  education_spending = w: republican
|  adoption_of_the_budget_resolution = y: democrat
|  adoption_of_the_budget_resolution = w: democrat
```

2.6 J48

The accuracy is 94.94% with default parameters.

```
Correctly Classified Instances         413               94.9425 %
Incorrectly Classified Instances        22                5.0575 %
```

The accuracy can be improved to 96.32% when we change the minNumObj from 2 to 7 and confidenceFactor from 0.25 to 1.

```
Correctly Classified Instances         419               96.3218 %
Incorrectly Classified Instances        16                3.6782 %
```

The classification rules are shown below:

```
J48 pruned tree
------------------

physician_fee_freeze = y
|   synfuels_corporation_cutback = w: republican (7.0)
|   synfuels_corporation_cutback = n: republican (138.0/3.0)
|   synfuels_corporation_cutback = y
|   |   adoption_of_the_budget_resolution = n: republican (23.0/5.0)
|   |   adoption_of_the_budget_resolution = y: democrat (8.0/3.0)
|   |   adoption_of_the_budget_resolution = w: democrat (1.0)
physician_fee_freeze = w: democrat (11.0/3.0)
physician_fee_freeze = n
|   adoption_of_the_budget_resolution = n
|   |   education_spending = y: democrat (10.0)
|   |   education_spending = n: democrat (14.0/1.0)
|   |   education_spending = w: republican (1.0)
|   adoption_of_the_budget_resolution = y: democrat (219.0)
|   adoption_of_the_budget_resolution = w: democrat (3.0)
```

2.7 LADTree

The accuracy is 94.94% with default parameters.

```
Correctly Classified Instances         413               94.9425 %
Incorrectly Classified Instances        22                5.0575 %
```

The accuracy can be increased to 96.32% when we change the numOfBoostingIterations from 10 to 20.

```
Correctly Classified Instances         419               96.3218 %
Incorrectly Classified Instances        16                3.6782 %
```

The classification rules is shown below(part):

```
: 0,0
|  (1)physician_fee_freeze = y: 0.977,-0.977
|  |  (6)immigration = n: -0.34,0.34
|  |  |  (9)mx_missile = n: 0.201,-0.201
|  |  |  |  (10)superfund_right_to_sue = y: 0.236,-0.236
|  |  |  |  |  (13)export_administration_act_sa = y: 0.571,-0.571
|  |  |  |  |  (13)export_administration_act_sa != y: -0.284,0.284
|  |  |  |  (10)superfund_right_to_sue != y: -0.987,0.987
|  |  |  |  |  (15)export_administration_act_sa = w: 0.588,-0.588
|  |  |  |  |  (15)export_administration_act_sa != w: -0.613,0.613
|  |  |  (9)mx_missile != n: -0.847,0.847
|  |  (6)immigration != n: 0.532,-0.532
|  (1)physician_fee_freeze != y: -1.236,1.236
|  (2)adoption_of_the_budget_resolution = y: -0.529,0.529
|  |  (7)nti_satellite_test_ban = y: 0.071,-0.071
|  |  |  (8)physician_fee_freeze = y: 1.403,-1.403
|  |  |  (8)physician_fee_freeze != y: -1.016,1.016
|  |  (7)nti_satellite_test_ban != y: -0.878,0.878
|  |  |  (11)water_project_cost_sharing = y: -0.679,0.679
|  |  |  (11)water_project_cost_sharing != y: 0.632,-0.632
|  (2)adoption_of_the_budget_resolution != y: 0.345,-0.345
|  (3)synfuels_corporation_cutback = y: -0.61,0.61
|  (3)synfuels_corporation_cutback != y: 0.217,-0.217
|  |  (4)physician_fee_freeze = n: -0.491,0.491
|  |  |  (12)superfund_right_to_sue = n: 0.573,-0.573
|  |  |  |  (16)adoption_of_the_budget_resolution = n: 0.716,-0.716
|  |  |  |  (16)adoption_of_the_budget_resolution != n: -0.512,0.512
|  |  |  (12)superfund_right_to_sue != n: -0.605,0.605
|  |  (4)physician_fee_freeze != n: 0.452,-0.452
|  |  (14)education_spending = n: -0.183,0.183
|  |  |  (17)religious_groups_in_schools = n: 0.465,-0.465
|  |  |  |  (18)crime = y: 0.837,-0.837
|  |  |  |  (18)crime != y: -0.786,0.786
|  |  |  (17)religious_groups_in_schools != n: -0.376,0.376
|  |  (14)education_spending != n: 0.352,-0.352
|  (5)education_spending = n: -0.435,0.435
|  (5)education_spending != n: 0.191,-0.191
```

2.8 LMT

We get the accuracy 96.78% using default parameters,

```
Correctly Classified Instances         421               96.7816 %
Incorrectly Classified Instances        14                3.2184 %
```

It makes no difference when we use different parameters.

The classification rules are shown below:

```
Logistic model tree
------------------
: LM_1:20/20 (435)

Number of Leaves  :     1

Size of the Tree :      1
LM_1:
Class 0 :
0.31 +
[water_project_cost_sharing=y] * -0.51 +
[adoption_of_the_budget_resolution=n] * 0.38 +
[adoption_of_the_budget_resolution=y] * -0.85 +
[physician_fee_freeze=y] * 1.8  +
[physician_fee_freeze=n] * -1.58 +
[nti_satellite_test_ban=n] * -0.38 +
[nti_satellite_test_ban=y] * 0.32 +
[aid_to_nicaraguan_contras=w] * 0.44 +
[mx_missile=y] * -0.54 +
[immigration=n] * -0.59 +
[synfuels_corporation_cutback=n] * 0.71 +
[synfuels_corporation_cutback=y] * -0.8 +
[education_spending=n] * -0.58 +
[education_spending=w] * 0.35 +
[superfund_right_to_sue=w] * -0.46 +
[duty_free_exports=y] * -0.41 +
[export_administration_act_sa=y] * 0.41

Class 1 :
-0.31 +
[water_project_cost_sharing=y] * 0.51 +
[adoption_of_the_budget_resolution=n] * -0.38 +
[adoption_of_the_budget_resolution=y] * 0.85 +
[physician_fee_freeze=y] * -1.8 +
[physician_fee_freeze=n] * 1.58 +
[nti_satellite_test_ban=n] * 0.38 +
[nti_satellite_test_ban=y] * -0.32 +
[aid_to_nicaraguan_contras=w] * -0.44 +
[mx_missile=y] * 0.54 +
[immigration=n] * 0.59 +
[synfuels_corporation_cutback=n] * -0.71 +
[synfuels_corporation_cutback=y] * 0.8  +
[education_spending=n] * 0.58 +
[education_spending=w] * -0.35 +
[superfund_right_to_sue=w] * 0.46 +
[duty_free_exports=y] * 0.41 +
[export_administration_act_sa=y] * -0.41
```

2.9 NBTree

The accuracy is 94.71% when we use default parameters.

```
Correctly Classified Instances         412               94.7126 %
Incorrectly Classified Instances        23                5.2874 %
```

The classification rules are shown below:

```
NBTree
------------------

nti_satellite_test_ban = n
|   aid_to_nicaraguan_contras = n: NB 2
|   aid_to_nicaraguan_contras = y: NB 3
|   aid_to_nicaraguan_contras = w: NB 4
nti_satellite_test_ban = y
|   superfund_right_to_sue = y: NB 6
|   superfund_right_to_sue = w: NB 7
|   superfund_right_to_sue = n
|   |   physician_fee_freeze = y: NB 9
|   |   physician_fee_freeze = w: NB 10
|   |   physician_fee_freeze = n
|   |   |   synfuels_corporation_cutback = w: NB 12
|   |   |   synfuels_corporation_cutback = n: NB 13
|   |   |   synfuels_corporation_cutback = y: NB 14
nti_satellite_test_ban = w: NB 15
```

2.10 RandomTree

The accuracy is 93.56% with default parameters.

```
Correctly Classified Instances         407               93.5632 %
Incorrectly Classified Instances        28                6.4368 %
```

The accuracy can be increased to 94.71% when we change the seed from 1 to 15.

```
Correctly Classified Instances         410               94.2529 %
Incorrectly Classified Instances        25                5.7471 %
```

The classification rules are shown below(part):

```
RandomTree
==========

aid_to_nicaraguan_contras = n
|   education_spending = y
|   |   synfuels_corporation_cutback = w : republican (3/0)
|   |   synfuels_corporation_cutback = n
|   |   |   physician_fee_freeze = y
|   |   |   |   adoption_of_the_budget_resolution = n : republican (90/0)
|   |   |   |   adoption_of_the_budget_resolution = y
|   |   |   |   |   water_project_cost_sharing = y
|   |   |   |   |   |   export_administration_act_sa = y : republican (3/0)
|   |   |   |   |   |   export_administration_act_sa = w : democrat (1/0)
|   |   |   |   |   |   export_administration_act_sa = n : republican (0/0)
|   |   |   |   |   water_project_cost_sharing = n : republican (3/0)
|   |   |   |   |   water_project_cost_sharing = w : republican (0/0)
|   |   |   |   adoption_of_the_budget_resolution = w : republican (1/0)
|   |   |   physician_fee_freeze = w : republican (0/0)
|   |   |   physician_fee_freeze = n : democrat (3/0)
|   |   synfuels_corporation_cutback = y
|   |   |   export_administration_act_sa = y
|   |   |   |   adoption_of_the_budget_resolution = n
|   |   |   |   |   superfund_right_to_sue = y
|   |   |   |   |   |   immigration = y
|   |   |   |   |   |   |   water_project_cost_sharing = y : republican (3/0)
|   |   |   |   |   |   |   water_project_cost_sharing = n : democrat (1/0)
|   |   |   |   |   |   |   water_project_cost_sharing = w : republican (0/0)
|   |   |   |   |   |   immigration = n : republican (4/0)
|   |   |   |   |   |   immigration = w : republican (0/0)
|   |   |   |   |   superfund_right_to_sue = w : republican (0/0)
|   |   |   |   |   superfund_right_to_sue = n
|   |   |   |   |   |   handicapped_infants = n : republican (1/0)
|   |   |   |   |   |   handicapped_infants = w : republican (0/0)
|   |   |   |   |   |   handicapped_infants = y : democrat (1/0)
```

2.11 SimpleCart

The accuracy is 95.63% with default parameters. This accuracy is not improved under different parameters.

```
Correctly Classified Instances         416               95.6322 %
Incorrectly Classified Instances        19                4.3678 %
```

The classification rules are shown below:

```
CART Decision Tree

physician_fee_freeze=(n)|(w)
|  adoption_of_the_budget_resolution=(y)|(n): democrat(247.0/2.0)
|  adoption_of_the_budget_resolution!=(y)|(n)
|  |  mx_missile=(n)|(y): democrat(6.0/1.0)
|  |  mx_missile!=(n)|(y): republican(2.0/0.0)
physician_fee_freeze!=(n)|(w)
|  synfuels_corporation_cutback=(y)
|  |  adoption_of_the_budget_resolution=(w)|(y)
|  |  |  nti_satellite_test_ban=(n)|(w): democrat(6.0/0.0)
|  |  |  nti_satellite_test_ban!=(n)|(w): republican(3.0/0.0)
|  |  adoption_of_the_budget_resolution!=(w)|(y)
|  |  |  el_salvador_aid=(n): democrat(2.0/0.0)
|  |  |  el_salvador_aid!=(n): republican(18.0/3.0)
|  synfuels_corporation_cutback!=(y): republican(142.0/3.0)
```

2.12 summary

According to all analysis above, it turns out that we get the best accuracy when we implement LMT and FT methods. Hence the highest accuracy is 96.78%

3 Observation and Conclusion

As it can be seen, LMT is the best method for the classification of Iris among all kinds of methods. As for the classification of Congressmen, LMT and FT are best methods. We obtain best accuracy in both classification using LMT method. In this way, LMT method has the best performance in this project.