

Introduction to Machine Learning Assignment 5

Name: Vivekanand Reddy Malipatel
CWID: A20524971

Question 1 Answer:

a) (The Python code for this is in the file: 1_a.py)

Output Screenshot:

```
Intro to ML/Assignments/Assignments/1a.py
Training Partition:
Group 0 (no claims): 11377
Group 1 (1 claim): 6575
Group 2 (2 claims): 2055
Group 3 (3 or more claims): 525

Testing Partition:
Group 0 (no claims): 3817
Group 1 (1 claim): 2282
Group 2 (2 claims): 687
Group 3 (3 or more claims): 195
```

Answer:

Training Partition:

Group 0 (no claims): 11377

Group 1 (1 claim): 6575

Group 2 (2 claims): 2055

Group 3 (3 or more claims): 525

Testing Partition:

Group 0 (no claims): 3817

Group 1 (1 claim): 2282

Group 2 (2 claims): 687

Group 3 (3 or more claims): 195

b) (The Python code for this is in the file: 1_b.py)

Output Screenshot:

```
Lowest AIC value on the Training partition 40514.30941979059
model producing the Lowest AIC value on the Training partition ['f_primary_age_tier', 'f_residence_location', 'f_fire_alarm_type', 'f_mile_fire_station', 'f_aoi_tier']
```

Answer:

Lowest AIC value on the Training partition 40514.30941979059

model producing the Lowest AIC value on the Training partition : ['f_primary_age_tier', 'f_residence_location', 'f_fire_alarm_type', 'f_mile_fire_station', 'f_aoi_tier']

c) (The Python code for this is in the file: 1_c.py)

Output Screenshot :

```
Lowest BIC value on the Training partition 40880.07276393123
model producing the Lowest BIC value on the Training partition ['f_primary_age_tier', 'f_fire_alarm_type', 'f_mile_fire_station', 'f_aoi_tier']
```

Answer :

Lowest BIC value on the Training partition 40880.07276393123
model producing the Lowest BIC value on the Training partition ['f_primary_age_tier',
'f_fire_alarm_type', 'f_mile_fire_station', 'f_aoi_tier']

d) (The Python code for this is in the file: 1_d.py)

Output Screenshot :

```
highest Accuracy value on the Testing partition 0.017619252256123763  
model producing the highest Accuracy value on the Testing partition ['f_primary_age_tier', 'f_primary_gender', 'f_marital', 'f_residence_location', 'f_fire_alarm_type', 'f_mile_fire_station', 'f_aoi_tier']
```

Answer :

highest Accuracy value on the Testing partition 0.017619252256123763
model producing the highest Accuracy value on the Testing partition ['f_primary_age_tier',
'f_primary_gender', 'f_marital', 'f_residence_location', 'f_fire_alarm_type', 'f_mile_fire_station',
'f_aoi_tier']

e) (The Python code for this is in the file: 1_d.py)

Output Screenshot :

```
lowest Root Average Squared Error on the Testing partition 1.7412889173079829  
model producing the lowest Root Average Squared Error on the Testing partition ['f_primary_age_tier', 'f_marital', 'f_residence_location', 'f_fire_alarm_type', 'f_mile_fire_station', 'f_aoi_tier']
```

Answer :

lowest Root Average Squared Error on the Testing partition 1.7412889173079829
model producing the lowest Root Average Squared Error on the Testing partition ['f_primary_age_tier',
'f_marital', 'f_residence_location', 'f_fire_alarm_type', 'f_mile_fire_station', 'f_aoi_tier']

Question 2 Answers:

a) (The Python code for this is in the file: 2_a.py)

Output Screenshot :

```
vivekmalipatel@Viveks-MacBook-Pro Assignment5 % /usr/local/bin/python3 "/Users/vivekmalipatel/Library/CloudStorage/GoogleDrive-vmalipatel@hawk.iit.edu/My Drive/Intro to ML/Assignments/Assignment5/2a.py"
The Misclassification Rate of the classification tree on the Training data at Iteration 0 : 0.2471959533758522
vivekmalipatel@Viveks-MacBook-Pro Assignment5 %
```

Answer :

The Misclassification Rate of the classification tree on the Training data at Iteration 0 :
0.2471959533758522

b) (The Python code for this is in the file: 2_b.py)

Output Screenshot :

```
.iit.edu/My Drive/Intro to ML/Assignments/Assignment5/2b.py"
Number of Iterations performed to achieve convergence : 19
vivekmalipatel@Viveks-MacBook-Pro Assignment5 %
```

Answer :

Number of Iterations performed to achieve convergence : 19

	Iteration	Sum of Weights	Weighted Accuracy
0	0	4547.0	0.7528040466241478
1	1	3298.1008263742933	0.8398342905040086
2	2	8602.232914649527	0.9755766326177174
3	3	3614.427040090972	0.9738398884068389
4	4	8975.477397381019	0.9934014157116959
5	5	3657.1996550126028	0.9871631849937752
6	6	8699.609749421743	0.9989365506261062
7	7	3823.272844236134	0.9922962645290412
8	8	9267.338476201143	0.9989697483749743
9	9	3709.49851588819	0.9995885640214832
10	10	8922.412080883547	0.9999379963581282
11	11	3831.0501974943654	0.9991376176867861
12	12	9505.01156145901	0.9998849333752113
13	13	3884.037597497886	0.999906352933896
14	14	9365.172683484956	0.9999908714219145
15	15	3876.058957357368	0.9999983667905822
16	16	9398.973974392971	0.9999997754987902
17	17	3960.000661252515	0.999999822381173
18	18	9431.997108003678	0.9999999751425136

c) (The Python code for this is in the file: 2_cde.py)

Output Screenshot :

```
vivekmalipatel@Viveks-MacBook-Pro Assignment5 % /usr/local/bin/python3 "/Users/v  
Intro to ML/Assignments/Assignment5/2cde.py"  
Area Under Curve on the Testing data using the final converged classification tree : 0.6902535164515995
```

Answer :

Area Under Curve on the Testing data using the final converged classification tree :
0.6902535164515995

d) (The Python code for this is in the file: 2_cde.py)

Output Screenshot :

```
Accuracy of the Testing data using the final converged classification tree : 0.5876923076923077
```

Answer :

Accuracy of the Testing data using the final converged classification tree :
0.5876923076923077

e) (The Python code for this is in the file: 2_cde.py)

Box Plot:

