Introduction to Machine Learning Assignment 5

Name: Vivekanand Reddy Malipatel

CWID: A20524971

**Question 1 Answer:**

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

Output Screenshot:



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

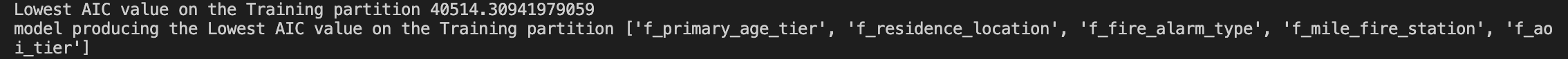
Group 1 (1 claim): 2282

Group 2 (2 claims): 687

Group 3 (3 or more claims): 195

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

Output Screenshot:



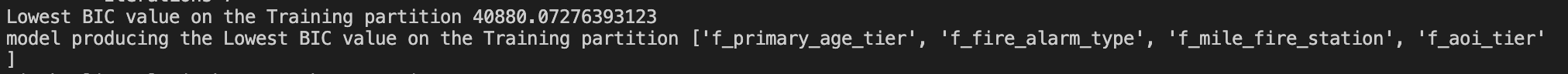
Answer:

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']

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

Output Screenshot :



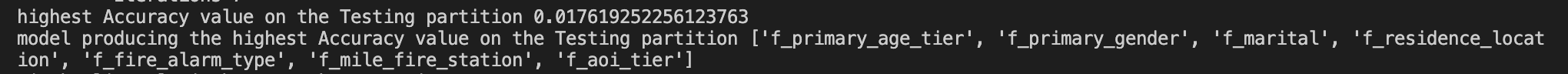
Answer :

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']

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

Output Screenshot :

****

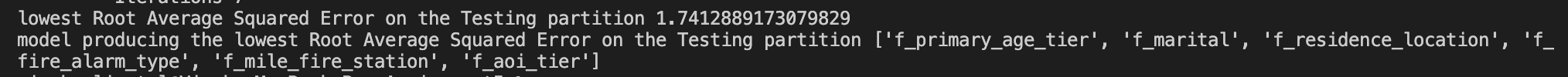
Answer :

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']

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

Output Screenshot :



Answer :

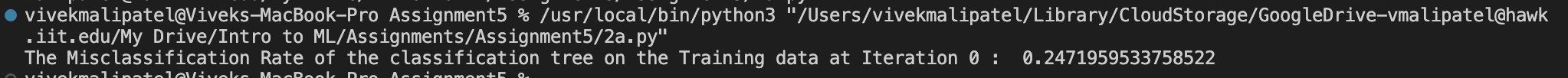
lowest Root Average Squared Error on the Testing partition 1.7412889173079829

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:**

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

Output Screenshot :

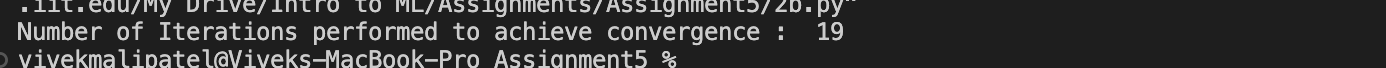


Answer :

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

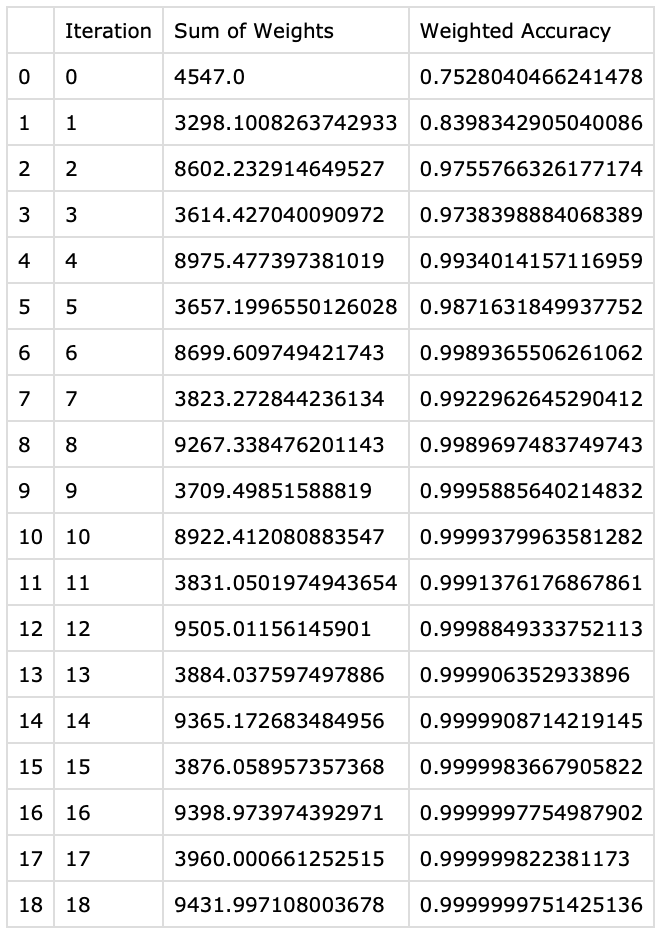
1. (The Python code for this is in the file: 2\_b.py)

Output Screenshot :



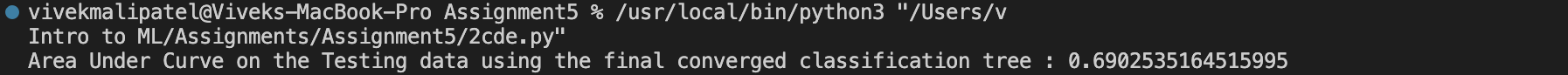
Answer :

Number of Iterations performed to achieve convergence : 19



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

Output Screenshot :



Answer :

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

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

Output Screenshot :



Answer :

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

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

Box Plot:

