3a. From the file “DT\_A.txt”, the output is:

att0 (IG:0.11667318106592728, Entropy:0.4021791902022729)

--att0 == 0--

att1 (IG:0.32181027983678356, Entropy:0.7320666900931938)

--att1 == 0--

leaf {0:0, 1:23}

--att1 == 1--

att2 (IG:1.0, Entropy:1.0)

--att2 == 0--

leaf {0:0, 1:8}

--att2 == 1--

leaf {0:8, 1:0}

--att0 == 1--

leaf {0:0, 1:61}

The accuracy as printed on terminal is 100%

3b. From the file “DT\_B.txt”, the output is:

att3 (IG:0.05319930896587699, Entropy:0.9542757225652247)

--att3 == 0--

att0 (IG:0.06470292426927621, Entropy:0.7896553247628193)

--att0 == 0--

att2 (IG:0.12296085056279982, Entropy:0.48977901368693755)

--att2 == 0--

leaf {0:0, 1:117}

--att2 == 1--

att1 (IG:0.3017760176219528, Entropy:0.7642045065086203)

--att1 == 0--

att4 (IG:0.9988455359952018, Entropy:0.9988455359952018)

--att4 == 0--

leaf {0:24, 1:0}

--att4 == 1--

leaf {0:0, 1:26}

--att1 == 1--

leaf {0:0, 1:58}

--att0 == 1--

att2 (IG:0.07402276613453679, Entropy:0.9383153522334069)

--att2 == 0--

att1 (IG:0.9967916319816366, Entropy:0.9967916319816366)

--att1 == 0--

leaf {0:0, 1:49}

--att1 == 1--

leaf {0:56, 1:0}

--att2 == 1--

att1 (IG:0.3057703826524583, Entropy:0.7670030768842513)

--att1 == 0--

att4 (IG:0.9993375041688847, Entropy:0.9993375041688847)

--att4 == 0--

leaf {0:32, 1:0}

--att4 == 1--

leaf {0:0, 1:34}

--att1 == 1--

leaf {0:0, 1:77}

--att3 == 1--

att4 (IG:0.16206771803574732, Entropy:0.9999766241093004)

--att4 == 0--

att2 (IG:0.00475352963095016, Entropy:0.8344210035612929)

--att2 == 0--

att1 (IG:0.4050728857879412, Entropy:0.8874179270685691)

--att1 == 0--

att0 (IG:0.9646900825612559, Entropy:0.9646900825612559)

--att0 == 0--

leaf {0:36, 1:0}

--att0 == 1--

leaf {0:0, 1:23}

--att1 == 1--

leaf {0:0, 1:59}

--att2 == 1--

att0 (IG:0.3241620856451307, Entropy:0.7829924501610913)

--att0 == 0--

leaf {0:0, 1:79}

--att0 == 1--

att1 (IG:0.9998393017810485, Entropy:0.9998393017810485)

--att1 == 0--

leaf {0:0, 1:33}

--att1 == 1--

leaf {0:34, 1:0}

--att4 == 1--

att2 (IG:0.3209981390018226, Entropy:0.8414100705725518)

--att2 == 0--

att0 (IG:9.483577640820862E-4, Entropy:0.9990389635262904)

--att0 == 0--

att1 (IG:1.0, Entropy:1.0)

--att1 == 0--

leaf {0:0, 1:34}

--att1 == 1--

leaf {0:34, 1:0}

--att0 == 1--

att1 (IG:0.9962088839046743, Entropy:0.9962088839046743)

--att1 == 0--

leaf {0:32, 1:0}

--att1 == 1--

leaf {0:0, 1:37}

--att2 == 1--

leaf {0:126, 1:0}

The accuracy as printed on terminal is 100%

3ci. The root feature will always have the maximum information gain. This is calculated by finding the entropy for each feature. Once that has been done, we find the information gain. The entropy with the least entropy will result in the maximum information gain which will be selected as the root node. For rtg\_A it is att0 and for rtg\_B it is att3.

3cii. Every feature in rtg\_B was utilized to construct the tree, however att3 in rtg\_A was not used for constructing the tree. Att3 only has 0s as its values. The entropy of that feature will equal the entropy of root. As entropy doesn’t change, information gain will be 0, so the feature will always be considered last.