Note: Lab- II Time: 2 hrs Write the program in Python.

1. Consider the following three tables of observations:

Id	Exercise	Family	Risk
1	daily	yes	low
2	weekly	yes	high
2	weekly	yes	high
5	rarely	no	high
5	rarely	no	high

Id	Smoker	Obese	Risk
1	false	false	low
2	true	false	high
2	true	false	high
4	true	true	high
5	true	true	high

Id	Obese	Family	Risk
1	false	yes	low
1	false	yes	low
2	false	yes	high
4	true	yes	high
5	true	no	high

The three tables above (separated by line) list three bootstrap samples, using these bootstrap samples create the decision trees that will be in the random forest model using entropy based information gain as the feature selection criteria. Assuming the random forest uses majority voting, what prediction will it return for the query: EXERCISE = rarely, SMOKER = false, OBESE = true, FAMILY = yes.

2. Consider a biometric matcher that generates similarity scores in the range $[0, 1]$. Its genuine and impostor score distributions are as follows: $p(s|genuine) = 4s^2$ and $p(s|impostor) = 4 - 4s^2$. Suppose the following decision rule is employed: s is classified as a genuine score. If $s \geq \eta$; else it is classified as an impostor score. Here, $\eta \in [0, 1]$.

3. Plot the genuine and impostor distributions in a single graph.

4. Write a program to compute the DET and ROC curves based on these two distributions. Plot the DET and ROC curves.

5. Consider a theoretical biometric matcher that generates distance scores in the range $[-\infty, +\infty]$. Assume that the genuine and impostor score distributions due to this matcher can be approximately modeled as $N(20, 5)$ and $N(60, 15)$, respectively. Here, $N(\mu, \sigma^2)$ denotes a normal distribution with mean, μ, and variance, σ^2. Suppose the following decision rule is employed: s is classified as

a genuine score if $s \leq \eta$; else it is classified as an impostor score. Here, $\eta \in [0, 100]$.

6. Plot the genuine and impostor distributions in a single graph. The distributions should be contained in the range $[0, 100]$.

7. Write a program to compute the DET and ROC curves based on these two distributions. Plot the DET and ROC curves.