

Programming Assignment 1
CSCI-860: Biometrics in a Networked Society

Presented To
Kiran Balagani
Associate Professor of Computer Science
New York Institute of Technology

Prepared By:
Pranav S. Krishnamurthy
Major – Computer Science
Student ID: 0704111

Due Date – 2nd December 2014
Submission Date – 4th December 2014

Problem Statement

This assignment requires the implementation of the Manhattan verifier to report *false accept* (*impostor pass*) and *false reject rates* on a publicly available keystroke biometric dataset. Any programming language may be used, as long as the program can be compiled on computers in Harry Schure Hall Room 212. In addition, a demonstration and explanation of our written code will be required.

My Approach

For this programming assignment, I have used the C#.net programming language, the Visual Studio 2013 Ultimate IDE for Windows 8.1 along with the PasswordData.csv file which I manipulated into 51 separate csv files, and when running the program, I have the code written such that depending upon the user number and the value of N, which represents the sample size: the training and testing vectors are extracted from the original data, the template vectors, genuine scores, impostor scores, and the various rates are calculated appropriately. The approach that I have taken for this assignment is detailed below:

1. Taking the CSV file from the link provided in the PDF file given, I took that file and broke that into 51 separate CSV files which represents each user.
2. Depending upon the user number that the end user enters (the user number will act as the “tag” for the subject ID), the program will then find the file in the respective directory, and convert the entire CSV file into a 2D double array.
3. When the end user enters a value for N, which can have the value of either 100, 200, or 300; the program will then parse through using the 2D double array, and the value of N and extract the testing samples which will be used to generate the mean vector.
4. The mean vector, the value of N, and the remaining samples are then parameters for generating the genuine scores. Here, I have assumed that the value of n in the formula for the Manhattan Distance is equal to the number of samples.
5. Then to generate the impostor scores, I will then be taking all the other user data which contain 400-N samples, and the mean vector of the user as parameters.
6. Finally, will be taking a look at the various genuine scores and using a threshold T and using the same threshold T, and looking at the various impostor scores, will calculate the Impostor Pass Rate and False Reject Rates. – This step, please note I will only be able have it such that I can find the Impostor Pass Rate which can also be referred to as the False Accept Rate. I will briefly talk about the concept of this measure

False Accept Rate

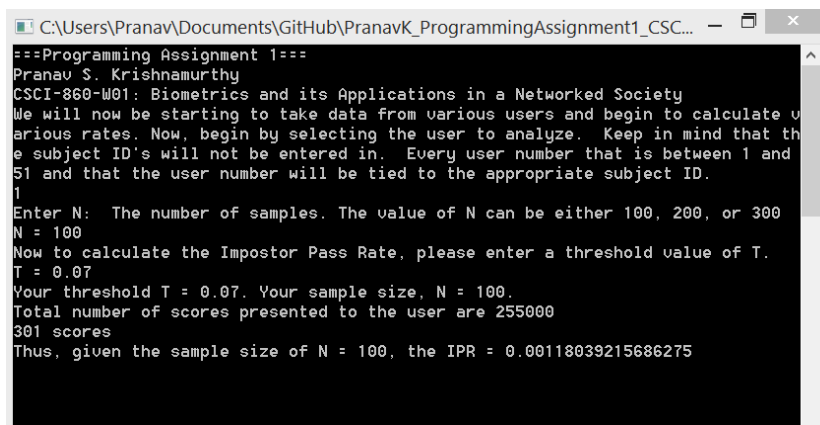
This is sometimes referred to as the impostor pass rate, to find this measure, I will be taking a threshold that the end user will put in and in all of the impostor scores that I have calculated, I will compare the value of the user entered threshold and the score. If the impostor score is less than or equal to the threshold value, that means the impostor has passed through the system. The screenshots below, I will show the values of thresholds that I have arbitrarily set. There will be three subsections which represent the various values of N, and also in each subset of screenshots, I will show the threshold values, and I will have five in total.

Value of $N = 100$ on User 1

In Figures 1 through 5 (below), you can see the following:

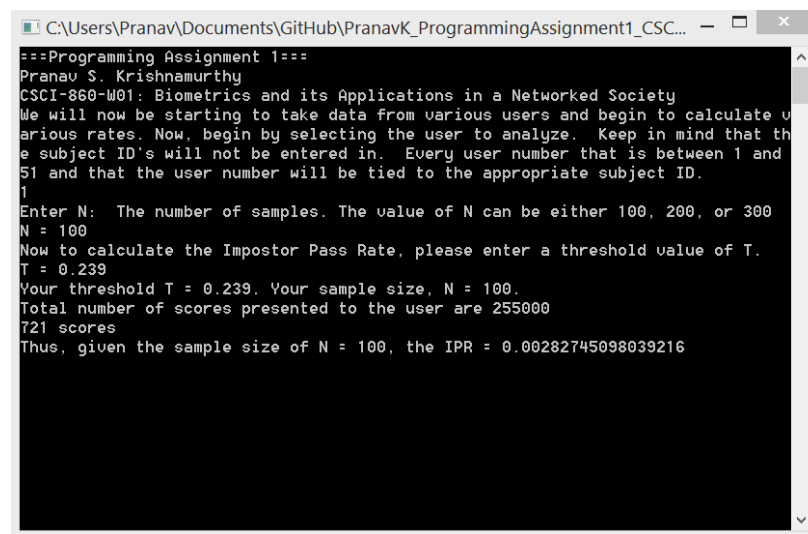
1. Threshold value of T being entered
2. Number of scores which satisfy the logic of “if the score is less than or equal to the threshold T , it is marked as genuine, else it is impostor”
3. Finally, you will be shown that for the number of samples, $N = 100$, you will be given the IPR.

Figure 1. The threshold value that I have set to be $T = 0.07$ and the Impostor Pass Rates are calculated, given the analysis of user 1 with the other users using the sample size, $N = 100$.



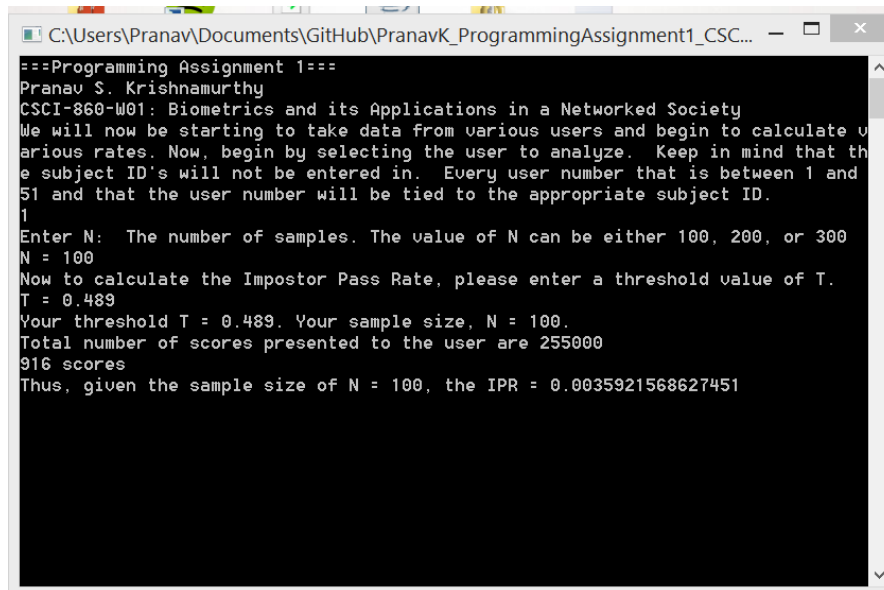
```
C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSC... - [icon] x
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate v
arious rates. Now, begin by selecting the user to analyze. Keep in mind that th
e subject ID's will not be entered in. Every user number that is between 1 and
51 and that the user number will be tied to the appropriate subject ID.
1
Enter N: The number of samples. The value of N can be either 100, 200, or 300
N = 100
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 0.07
Your threshold T = 0.07. Your sample size, N = 100.
Total number of scores presented to the user are 255000
301 scores
Thus, given the sample size of N = 100, the IPR = 0.00118039215686275
```

Figure 2. The threshold value that I have set to be $T = 0.239$ and the Impostor Pass Rates are calculated, given the analysis of user 1 with the other users using the sample size, $N = 100$.



```
C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSC... - [icon] x
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate v
arious rates. Now, begin by selecting the user to analyze. Keep in mind that th
e subject ID's will not be entered in. Every user number that is between 1 and
51 and that the user number will be tied to the appropriate subject ID.
1
Enter N: The number of samples. The value of N can be either 100, 200, or 300
N = 100
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 0.239
Your threshold T = 0.239. Your sample size, N = 100.
Total number of scores presented to the user are 255000
721 scores
Thus, given the sample size of N = 100, the IPR = 0.00282745098039216
```

Figure 3. The threshold value that I have set to be $T = 0.489$ and the Impostor Pass Rates are calculated given the analysis of user 1 with the other users, using the sample size $N = 100$.

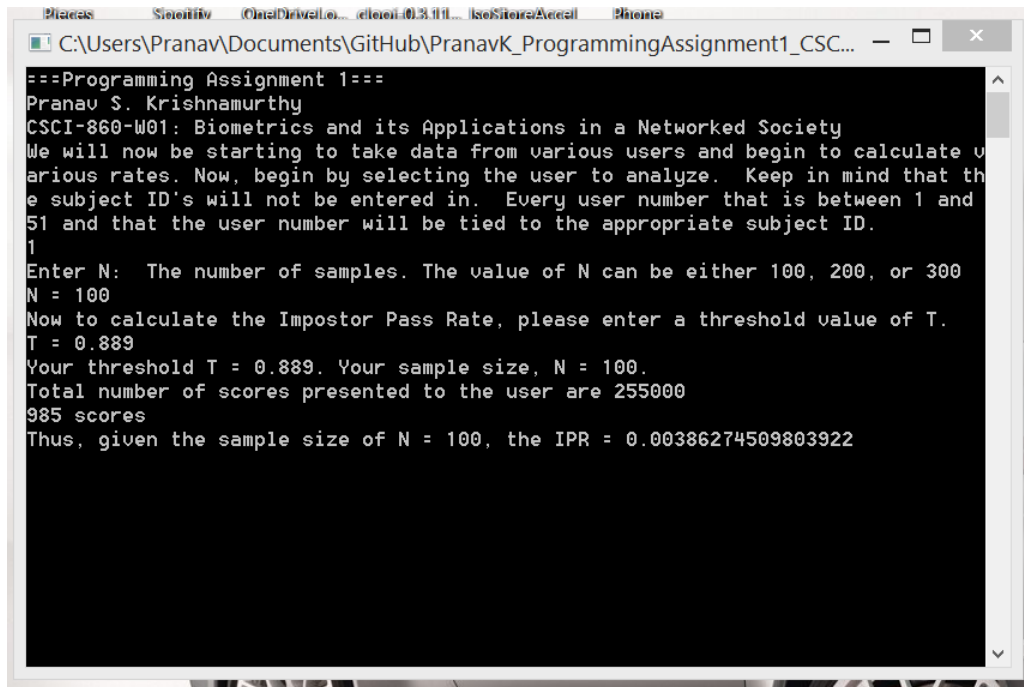


```

C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSC...
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate v
various rates. Now, begin by selecting the user to analyze. Keep in mind that th
e subject ID's will not be entered in. Every user number that is between 1 and
51 and that the user number will be tied to the appropriate subject ID.
1
Enter N: The number of samples. The value of N can be either 100, 200, or 300
N = 100
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 0.489
Your threshold T = 0.489. Your sample size, N = 100.
Total number of scores presented to the user are 255000
916 scores
Thus, given the sample size of N = 100, the IPR = 0.0035921568627451

```

Figure 4. The threshold value I have set to be $T = 0.889$ and the Impostor Pass Rates are calculated given the analysis of user 1 with the other users, using the sample size $N = 100$.

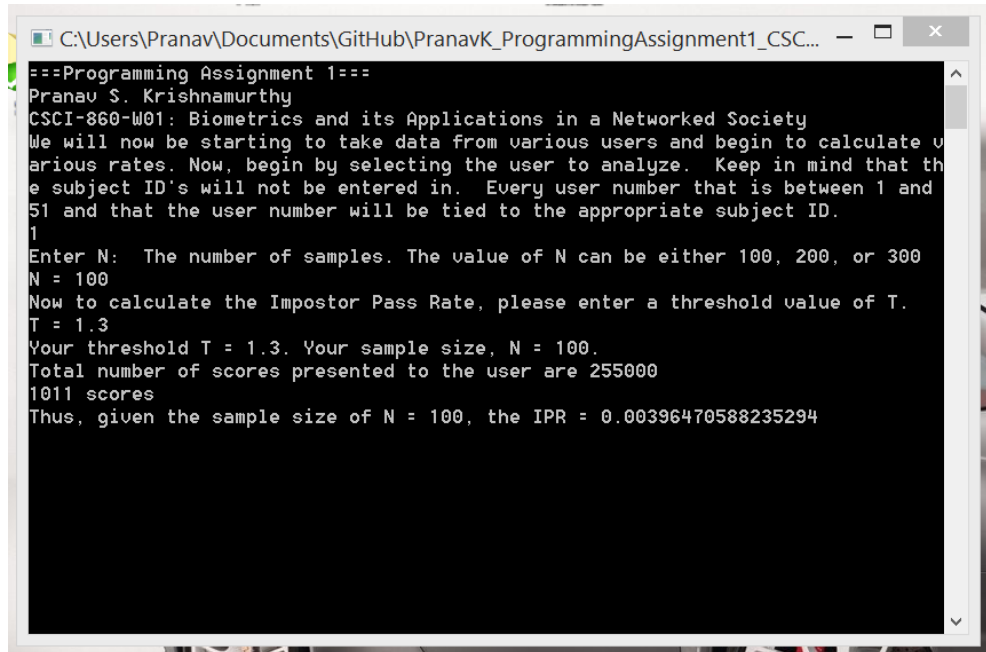


```

C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSC...
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate v
various rates. Now, begin by selecting the user to analyze. Keep in mind that th
e subject ID's will not be entered in. Every user number that is between 1 and
51 and that the user number will be tied to the appropriate subject ID.
1
Enter N: The number of samples. The value of N can be either 100, 200, or 300
N = 100
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 0.889
Your threshold T = 0.889. Your sample size, N = 100.
Total number of scores presented to the user are 255000
985 scores
Thus, given the sample size of N = 100, the IPR = 0.00386274509803922

```

Figure 5. The threshold value I have set to be $T = 1.3$ and the Impostor Pass Rates are calculated given the analysis of user 1 with the other users, using the sample size $N = 100$.

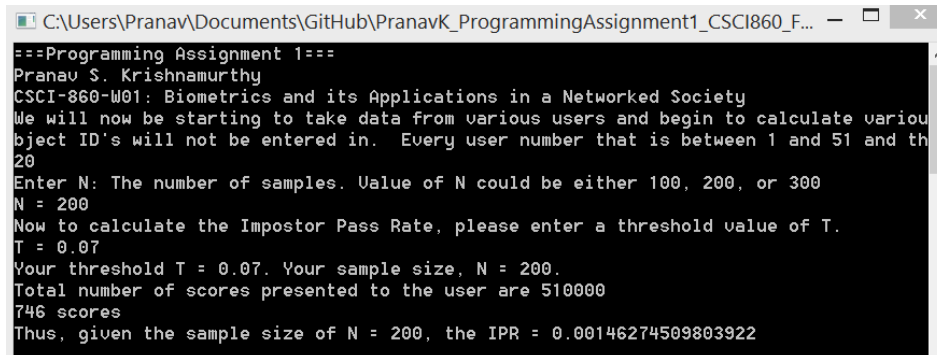


```
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate various rates. Now, begin by selecting the user to analyze. Keep in mind that the subject ID's will not be entered in. Every user number that is between 1 and 51 and that the user number will be tied to the appropriate subject ID.
1
Enter N: The number of samples. The value of N can be either 100, 200, or 300
N = 100
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 1.3
Your threshold T = 1.3. Your sample size, N = 100.
Total number of scores presented to the user are 255000
1011 scores
Thus, given the sample size of N = 100, the IPR = 0.00396470588235294
```

Value of $N = 200$ on User 20

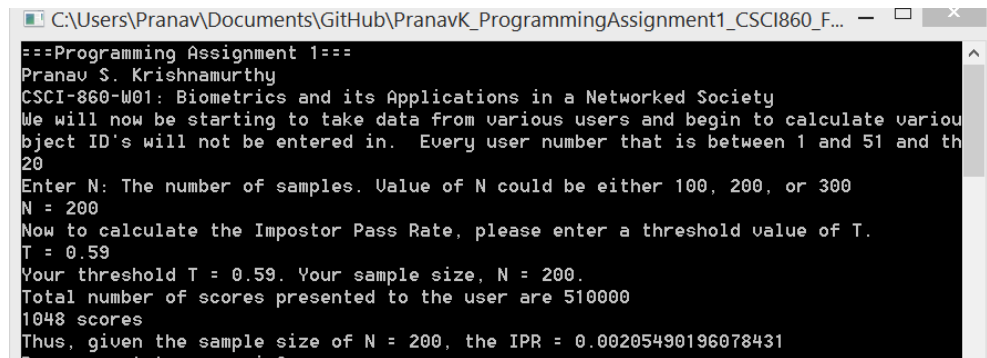
In the following figures, Figure 6 through 10 (below) I am going to be reporting the same information from the first group of screenshots above.

Figure 6. Given the sample size $N = 200$, $T = 0.07$, and the user number as 20. The Impostor Pass Rates are calculated accordingly.



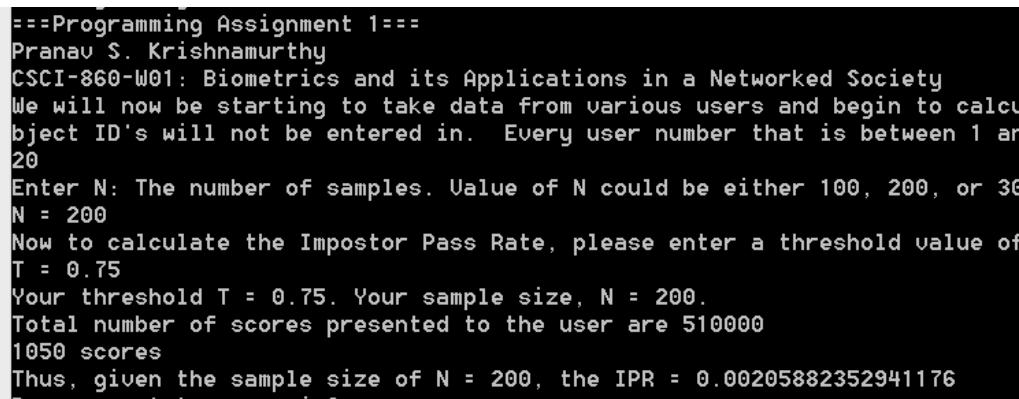
```
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate various rates. Now, begin by selecting the user to analyze. Keep in mind that the subject ID's will not be entered in. Every user number that is between 1 and 51 and that the user number will be tied to the appropriate subject ID.
20
Enter N: The number of samples. Value of N could be either 100, 200, or 300
N = 200
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 0.07
Your threshold T = 0.07. Your sample size, N = 200.
Total number of scores presented to the user are 510000
746 scores
Thus, given the sample size of N = 200, the IPR = 0.00146274509803922
```

Figure 7. Given the sample size $N = 200$, $T = 0.59$, and the user number as 20. The Impostor Pass Rates are calculated accordingly.



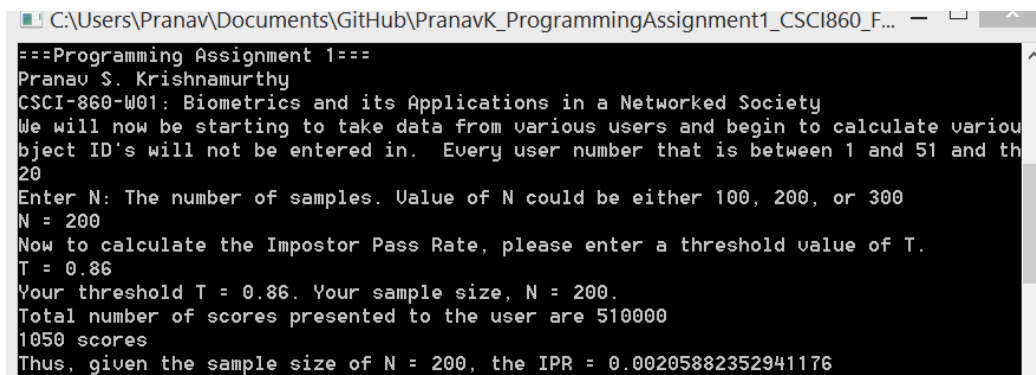
```
C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSCI860_F...  
===Programming Assignment 1===  
Pranav S. Krishnamurthy  
CSCI-860-W01: Biometrics and its Applications in a Networked Society  
We will now be starting to take data from various users and begin to calculate variou  
bject ID's will not be entered in. Every user number that is between 1 and 51 and th  
20  
Enter N: The number of samples. Value of N could be either 100, 200, or 300  
N = 200  
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.  
T = 0.59  
Your threshold T = 0.59. Your sample size, N = 200.  
Total number of scores presented to the user are 510000  
1048 scores  
Thus, given the sample size of N = 200, the IPR = 0.00205490196078431
```

Figure 8. Given the sample size $N = 200$, $T = 0.75$, and the user number as 20. The Impostor Pass Rates are calculated accordingly.



```
===Programming Assignment 1===  
Pranav S. Krishnamurthy  
CSCI-860-W01: Biometrics and its Applications in a Networked Society  
We will now be starting to take data from various users and begin to calcul  
bject ID's will not be entered in. Every user number that is between 1 and  
20  
Enter N: The number of samples. Value of N could be either 100, 200, or 300  
N = 200  
Now to calculate the Impostor Pass Rate, please enter a threshold value of  
T = 0.75  
Your threshold T = 0.75. Your sample size, N = 200.  
Total number of scores presented to the user are 510000  
1050 scores  
Thus, given the sample size of N = 200, the IPR = 0.00205882352941176
```

Figure 9. Given the sample size $N = 200$, $T = 0.86$, and the user number as 20. The Impostor Pass Rates are calculated accordingly.



```
C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSCI860_F...  
===Programming Assignment 1===  
Pranav S. Krishnamurthy  
CSCI-860-W01: Biometrics and its Applications in a Networked Society  
We will now be starting to take data from various users and begin to calculate variou  
bject ID's will not be entered in. Every user number that is between 1 and 51 and th  
20  
Enter N: The number of samples. Value of N could be either 100, 200, or 300  
N = 200  
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.  
T = 0.86  
Your threshold T = 0.86. Your sample size, N = 200.  
Total number of scores presented to the user are 510000  
1050 scores  
Thus, given the sample size of N = 200, the IPR = 0.00205882352941176
```

Figure 10. Given the sample size $N = 200$, $T = 0.94$, and the user number 20. The Impostor Pass Rates are calculated accordingly.

```

C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSCI860_F...
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate variou
bject ID's will not be entered in. Every user number that is between 1 and 51 and th
20
Enter N: The number of samples. Value of N could be either 100, 200, or 300
N = 200
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 0.94
Your threshold T = 0.94. Your sample size, N = 200.
Total number of scores presented to the user are 510000
1050 scores
Thus, given the sample size of N = 200, the IPR = 0.00205882352941176

```

Value of $N = 300$ for User 40

In the following figures, Figures 11 through 15, I will be conducting the same process where I will be randomly selecting various thresholds.

Figure 11. Sample size, $N = 300$, $T = 0.12$, and user number is 40. The Impostor Pass Rate is appropriately calculated.

```

C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSCI860_F...
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate variou
bject ID's will not be entered in. Every user number that is between 1 and 51 and th
40
Enter N: The number of samples. Value of N could be either 100, 200, or 300
N = 300
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 0.12
Your threshold T = 0.12. Your sample size, N = 300.
Total number of scores presented to the user are 765000
1046 scores
Thus, given the sample size of N = 300, the IPR = 0.00136732026143791

```

Figure 12. Sample size, $N = 300$, $T = 0.65$, and user number is 40. The Impostor Pass Rate is appropriately calculated.

```

C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSCI860_F...
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate variou
bject ID's will not be entered in. Every user number that is between 1 and 51 and th
40
Enter N: The number of samples. Value of N could be either 100, 200, or 300
N = 300
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 0.64
Your threshold T = 0.64. Your sample size, N = 300.
Total number of scores presented to the user are 765000
1050 scores
Thus, given the sample size of N = 300, the IPR = 0.00137254901960784

```

Figure 13. Sample size, $N = 300$, $T = 0.75$, user number is 40. The Impostor Pass Rate is appropriately calculated.

```
C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSCI860_F... - [X]
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate various
object ID's will not be entered in. Every user number that is between 1 and 51 and the
40
Enter N: The number of samples. Value of N could be either 100, 200, or 300
N = 300
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 0.75
Your threshold T = 0.75. Your sample size, N = 300.
Total number of scores presented to the user are 765000
1050 scores
Thus, given the sample size of N = 300, the IPR = 0.00137254901960784
```

Figure 14. Sample size, $N = 300$, user number is 40. The Impostor Pass Rate is appropriately calculated.

```
C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSCI860_F... - [X]
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate various
object ID's will not be entered in. Every user number that is between 1 and 51 and the
40
Enter N: The number of samples. Value of N could be either 100, 200, or 300
N = 300
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 0.9
Your threshold T = 0.9. Your sample size, N = 300.
Total number of scores presented to the user are 765000
1050 scores
Thus, given the sample size of N = 300, the IPR = 0.00137254901960784
```

Figure 15. Sample size, $N = 300$, $T = 1.67$, and user number is 40. The Impostor Pass Rate is appropriately calculated.

```
C:\Users\Pranav\Documents\GitHub\PranavK_ProgrammingAssignment1_CSCI860_F... - [X]
===Programming Assignment 1===
Pranav S. Krishnamurthy
CSCI-860-W01: Biometrics and its Applications in a Networked Society
We will now be starting to take data from various users and begin to calculate various
object ID's will not be entered in. Every user number that is between 1 and 51 and the
40
Enter N: The number of samples. Value of N could be either 100, 200, or 300
N = 300
Now to calculate the Impostor Pass Rate, please enter a threshold value of T.
T = 1.67
Your threshold T = 1.67. Your sample size, N = 300.
Total number of scores presented to the user are 765000
1050 scores
Thus, given the sample size of N = 300, the IPR = 0.00137254901960784
Do you want to go again?
```