

Authentication Experiments – Ideal Conditions/Weak Enrollment

- These experiments use the 36-subject DeskFree and LapFree data. This data is located on the 'utopia' server in <http://utopia.csis.pace.edu/cs691/2008-2009/team4/RawData36.zip>. You can also go to <http://utopia.csis.pace.edu/cs691/2008-2009/team4/> and select the Raw Data Used for Experimentation link under 'Other Resources' on the menu of choices on the left of the page. Download this zip file to your local computer.
- You will use the Refactored-BAS.zip file for all experiments. Be sure to use this version. You can replace the version you used for the prior common deliverables or use a new directory. Save the demographics.txt file if you replace previous software. You will use this file again. The required software can be downloaded from utopia: <http://utopia.csis.pace.edu/cs691/2009-2010/team4/team3/deliverables/final/Refactored-BAS.zip>. Alternately, you can download the file from <http://utopia.csis.pace.edu/cs691/2009-2010/team4/team3/> from the Deliverables link on the far left of the far left of the page. Scroll down to the bottom of the deliverables page and select Refactored-BAS.zip from the Final Refactored BAS System section.
- Using the techniques you learned in the first three common deliverables, process the raw data using the Bio Feature Extractor. After unzipping the raw data, you will see four directories, FableDesktop, FableLaptop, FreeDesktop, and FreeLaptop. For the following experiments we will only use the FreeDesktop and FreeLaptop data.

Create Working Directories

You can use any directories of your choosing. The examples in this guide will use the following directory structures:

Raw Data:

C:\it691\Keystroke A Experiments\FreeDesktop

C:\it691\Keystroke A Experiments\FreeLaptop

Demographics File Location:

C:\it691\demographics\demographics.txt

Feature Extractor Output:

C:\it691\keystroke A Experiments\FreeDesktopFeatures\freedesktopfeatures.txt

C:\it691\Keystroke A experiments\FreeLaptopFeatures\freelaptopfeatures.txt

Test Data

C:\it691\keystroke A Experiments\FreeDesktopTestData\bot18-test.txt

C:\it691\keystroke A Experiments\FreeLaptopTestData\bot18-test.txt

Train Data

C:\it691\keystroke A Experiments\FreeDesktopTrainData\top18-train.txt

C:\it691\keystroke A Experiments\FreeLaptopTrainData\topt18-train.txt

FreeDesktop Results

C:\it691\keystroke A Experiments\FreeDesktopResults

FreeLaptop Results

C:\it691\keystroke A Experiments\FreeLaptopResults

- **Verify the FreeDesktop Data**

This data set contains 176 text files. If you do not have 176 text files in the FreeDesktop data directory, verify that you extracted the correct data file and verify you are looking in the FreeDesktop directory. This file will not divide equally into testing and training files after feature extraction.

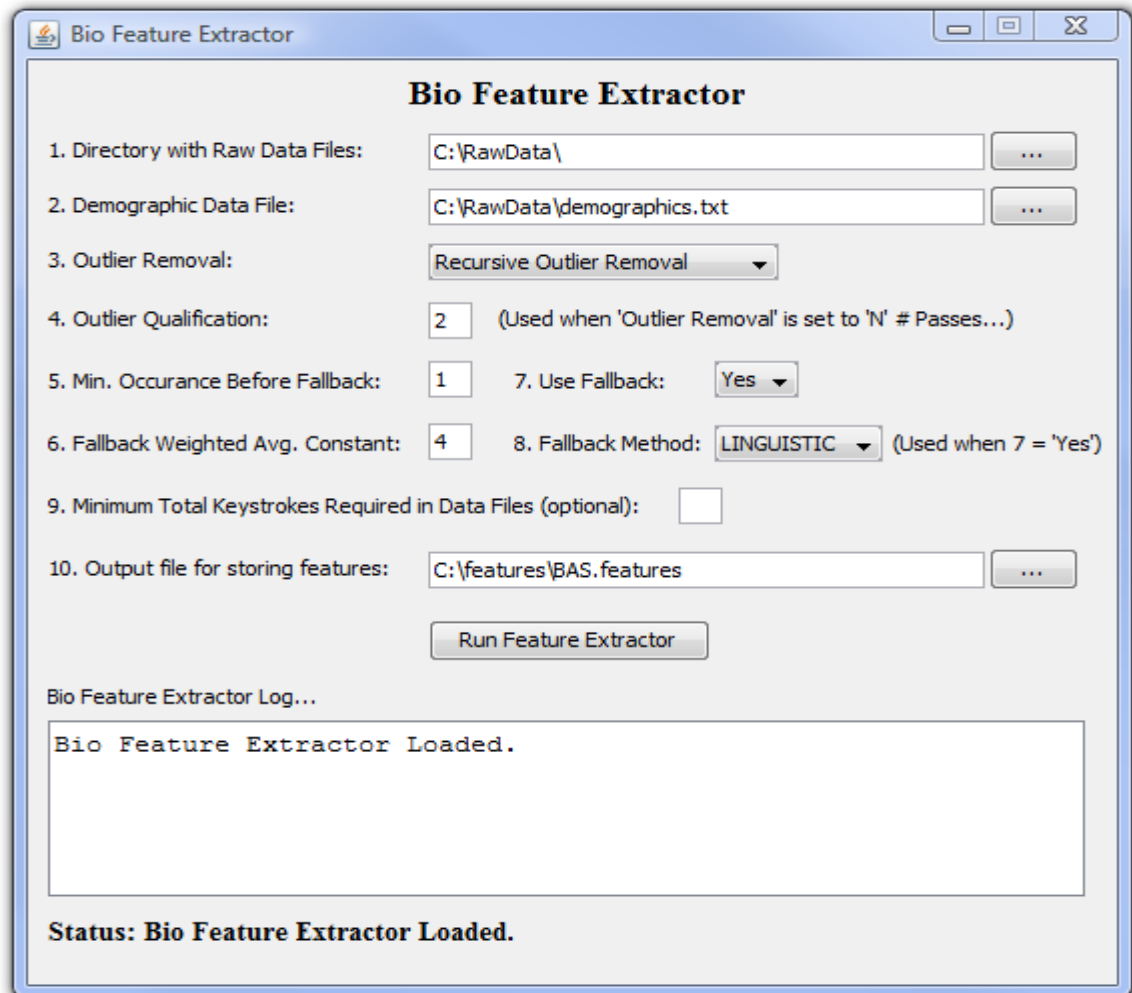
- **Verify the FreeLaptop Data**

This data set contains 180 text files. If you do not have 180 text files in the FreeLaptop data directory, verify that you extracted the correct data file and verify you are looking in the FreeLaptop directory. This file will divide equally into testing and training files after feature extraction.

- Bio Feature Extractor

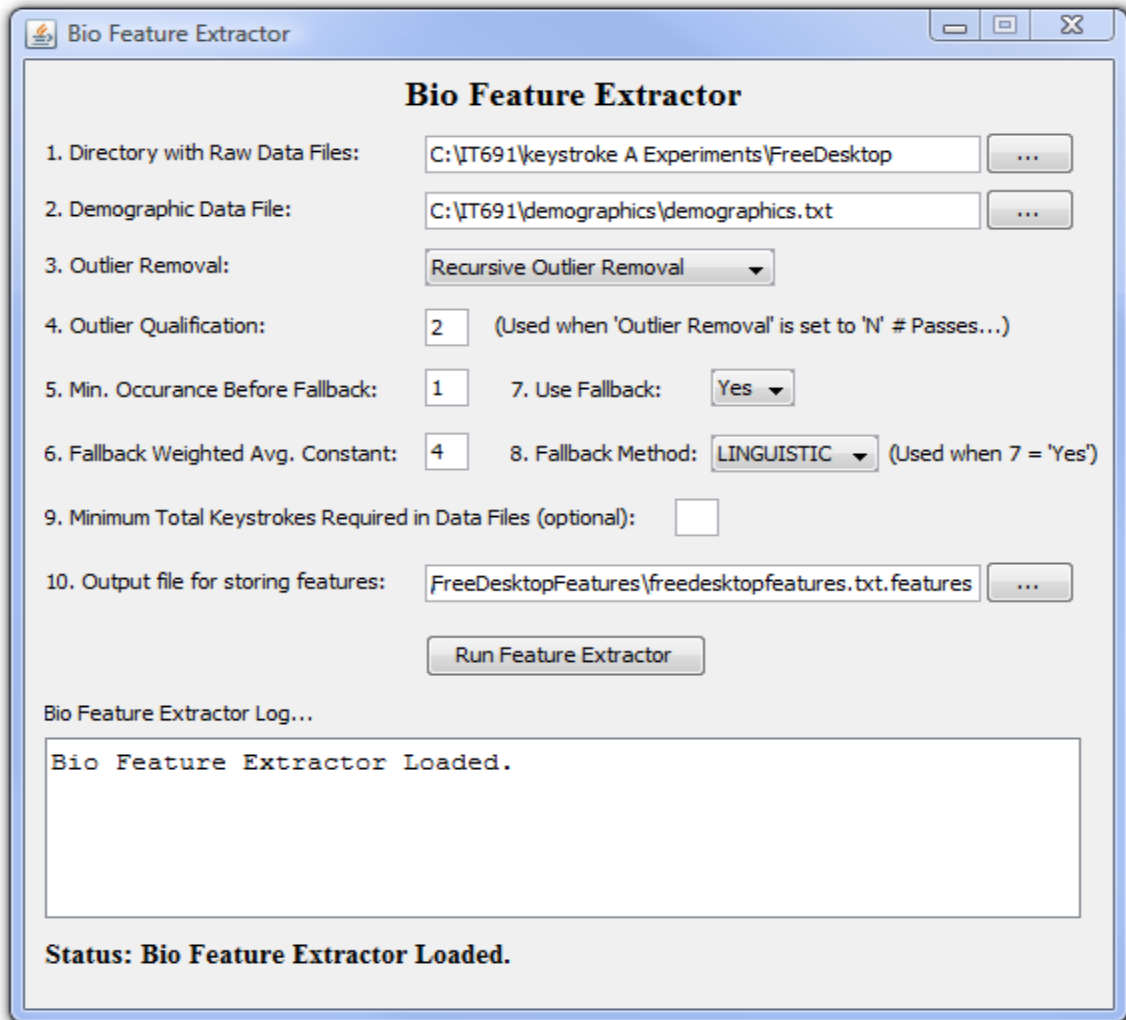
Launch the Bio Feature Extractor

Be sure to set the Raw Data File directory, Demographic Data File directory, and feature output directory to match your choices for these locations. If the Bio Feature Extractor does not appear similar to the following screen, you are likely using an older version.



Change Bio Feature Extractor settings for 1., 2., and 10. to match your specifications.

- **Desktop Free text Data**



Run Feature Extractor (biofeature.bat)

Press the 'Run Feature Extractor' button near the middle of the screen. Processing should only take a few seconds.

The bottom part of the on-screen log should display output similar to:

Success!!! New Features file created (176 files examined).

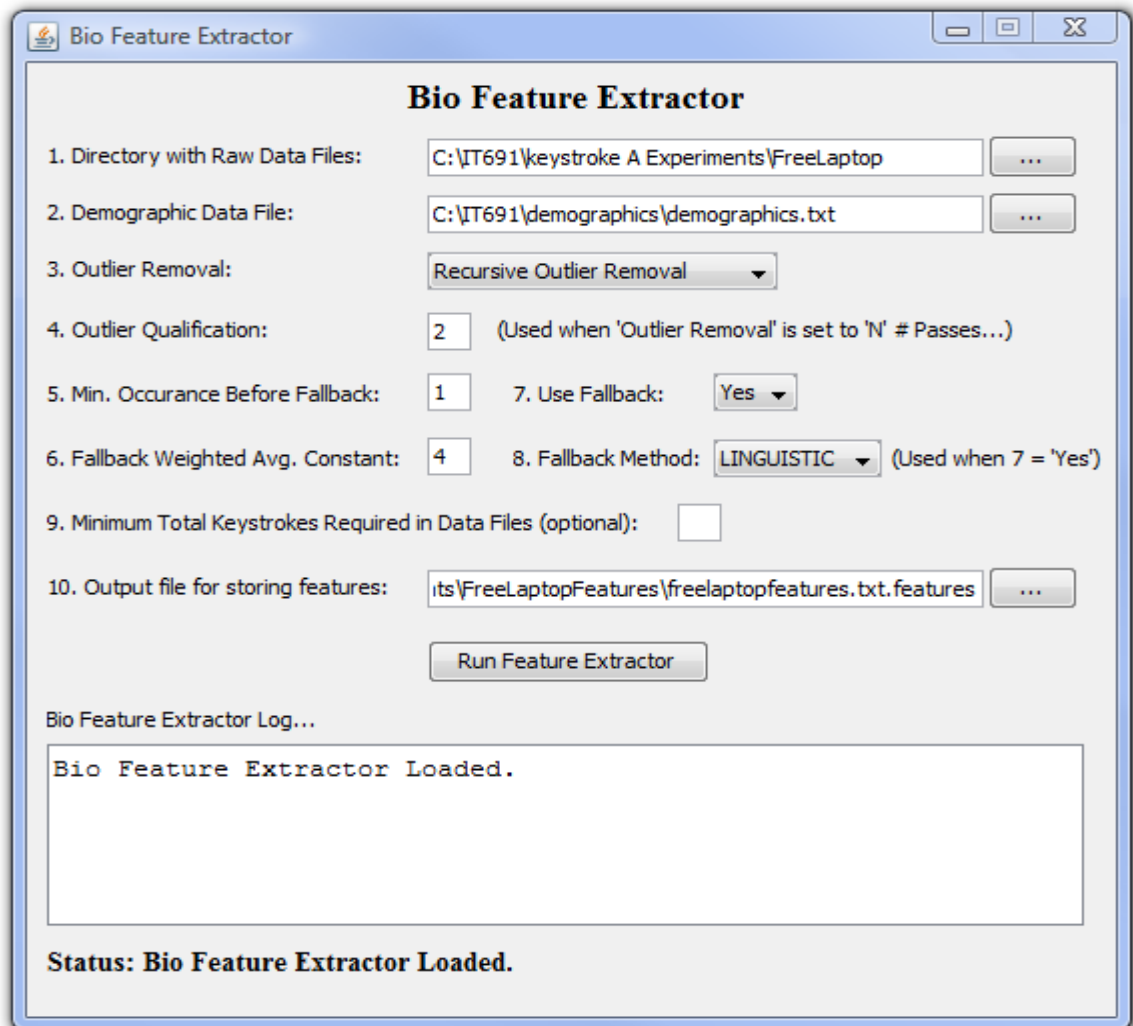
Features file location:

C:\IT691\keystroke A Experiments\FreeDesktopFeatures\freedesktopfeatures.txt.features

Feature extraction finished at Tue Feb 09 21:49:42 EST 2010.

Feature extraction completed in 4 seconds.

- **Laptop Free text Data**



Run Feature Extractor

Press the 'Run Feature Extractor' button near the middle of the screen. Processing should only take a few seconds.

The bottom part of the on-screen log should display output similar to:

Success!!! New Features file created (180 files examined).

Features file location:

C:\IT691\keystroke A Experiments\FreeLaptopFeatures\freelaptopfeatures.txt.features


Feature extraction finished at Tue Feb 09 21:59:23 EST 2010.

Feature extraction completed in 5 seconds.


- Verify Feature Files

In your designated feature file output directories, you should have the following files:

C:\it691\keystroke A Experiments\FreeDesktopFeatures\freedesktopfeatures.txt.features

Name	Date modified	Type	Size
 freedesktopfeatures.txt.features	2/9/2010 9:49 PM	FEATURES File	805 KB

C:\it691\keystroke A Experiments\FreeLaptopFeatures\freelaptopfeatures.txt.features

Name	Date modified	Type	Size
 freelaptopfeatures.txt.features	2/9/2010 9:59 PM	FEATURES File	832 KB

- Close the Bio Feature Extractor it by clicking on the 'X' in the top right of the dialogue box.
- Prepare the features file for BAS

DesktopFree

a) Create Training File

Training Data is comprised of data samples for the top 18 subjects. The resulting file should contain 89 rows and a two row header. Be sure to adjust the header to reflect the correct new row count.

```
Keystroke biometric data, Desktop/Free, created 2/9/2010
89
ANCEY VARGHESE/Female/20 - 29, Right, DESKTOP, FREETEXT1 1, 230, 0.67
ANCEY VARGHESE/Female/20 - 29, Right, DESKTOP, FREETEXT1 1, 230, 0.67
ANCEY VARGHESE/Female/20 - 29, Right, DESKTOP, FREETEXT1 1, 230, 0.67
ANCEY VARGHESE/Female/20 - 29, Right, DESKTOP, FREETEXT1 4, 230, 0.67
ANCEY VARGHESE/Female/20 - 29, Right, DESKTOP, FREETEXT1 4, 230, 0.67
ANNIE GRENCI/Female/30 - 39, Right, DESKTOP, FREETEXT1 1, 230, 0.67
ANNIE GRENCI/Female/30 - 39, Right, DESKTOP, FREETEXT1 1, 230, 0.59
```

Save the training file in the directory you created in Step 3 above.

b) Create Testing File

Testing Data is comprised of data samples for the bottom 18 subjects. You will need to remove the following two extraneous rows from the data as there is an invalid subject in the data. Remove the two rows for the subject FIRSTNAME LASTNAME.

```

112 FIRSTNAME LASTNAME/??/? ? DESKTOP FREETEXT8 1, 230, 0.7
113 FIRSTNAME LASTNAME/??/? ? DESKTOP FREETEXT3 1, 230, 0.8

```

Save the testing file in the directory you created in Step 3 above.

```

Keystroke biometric data, Desktop/Free, created 2/9/2010
87
KATHLEEN VILLANI/Female/50 - 59, Right, DESKTOP, FREETEXT7 1, 230, 0.673
KATHLEEN VILLANI/Female/50 - 59, Right, DESKTOP, FREETEXT7 1, 230, 0.760
KATHLEEN VILLANI/Female/50 - 59, Right, DESKTOP, FREETEXT7 1, 230, 0.654
KATHLEEN VILLANI/Female/50 - 59, Right, DESKTOP, FREETEXT7 1, 230, 0.695
KATHLEEN VILLANI/Female/50 - 59, Right, DESKTOP, FREETEXT7 1, 230, 0.623
LAUREN BOYLAN/Female/40 - 49, Right, DESKTOP, FABLE 1, 230, 0.9489859914
LAUREN BOYLAN/Female/40 - 49, Right, DESKTOP, FABLE 1, 230, 1.0, 0.95388
LAUREN BOYLAN/Female/40 - 49, Right, DESKTOP, FABLE 1, 230, 0.8903834599
LAUREN BOYLAN/Female/40 - 49, Right, DESKTOP, FABLE 4, 230, 0.9866947090

```

LaptopFree

a) Create Training File

Training Data is comprised of data samples for the top 18 subjects. The resulting file should contain 90 rows and a two row header. Be sure to adjust the header to reflect the correct new row count.

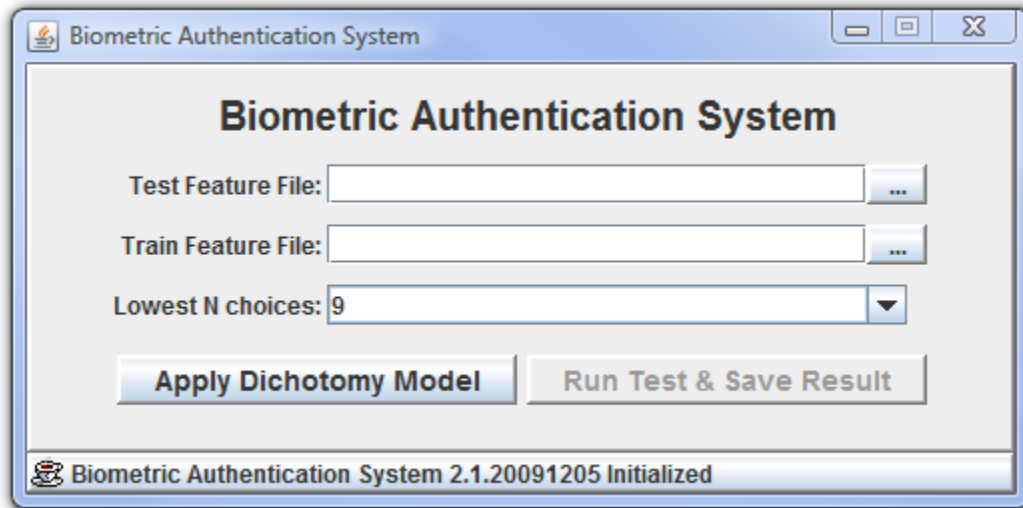
Save the training file in the directory you created in Step 3 above.

b) Create Testing File

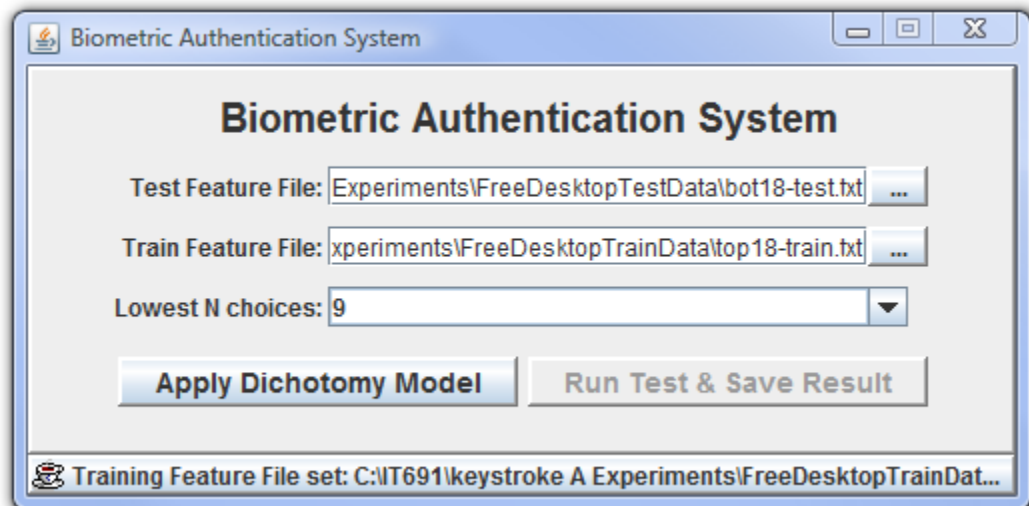
Testing Data is comprised of data samples for the bottom 18 subjects. The resulting file should contain 90 rows and a two row header. Be sure to adjust the header to reflect the correct new row count.

Save the training file in the directory you created in Step 3 above.

- Run BAS (Refactored-BAS.bat)

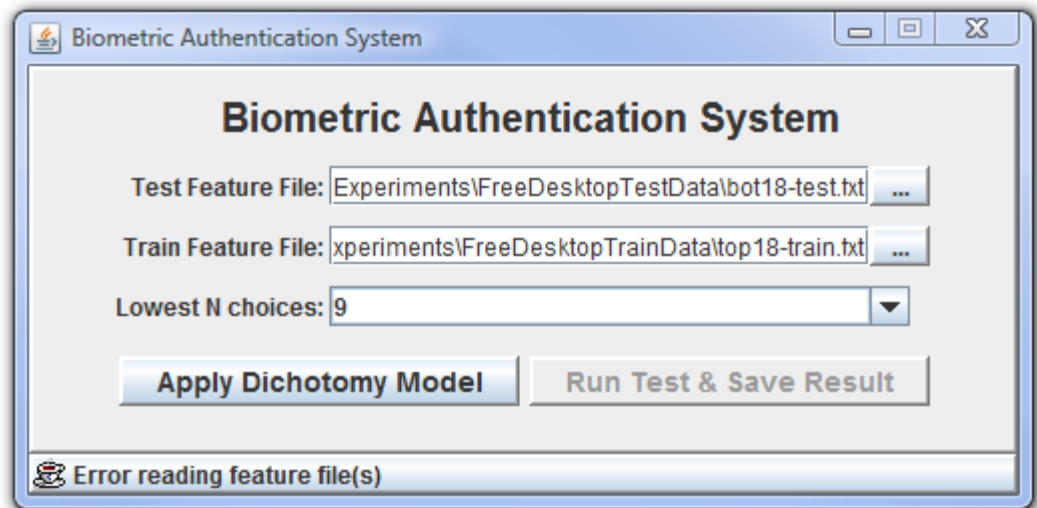


- Desktop Free Experiments
 - a) Select the Test and Train Feature Files for Desktop Free data that you created in the previous step.
 - b) Select the Lowest N choice of 9 (default value).

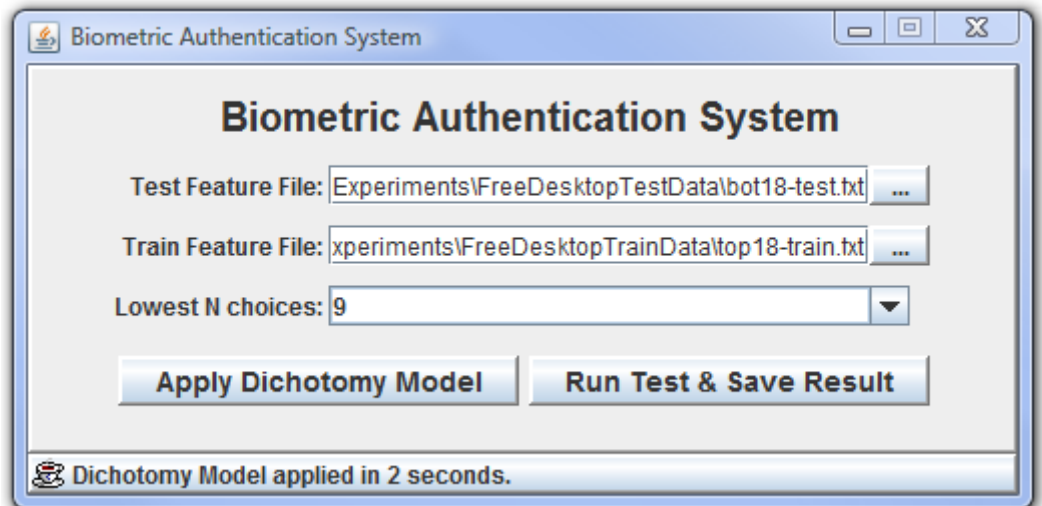


- c) Apply Dichotomy Model
- d) This experiment will use all dichotomy data. Select 0 for the maximum amount of dichotomy data for no maximum ceiling.

- e) If your test/train file headers are incorrect you will receive an error:

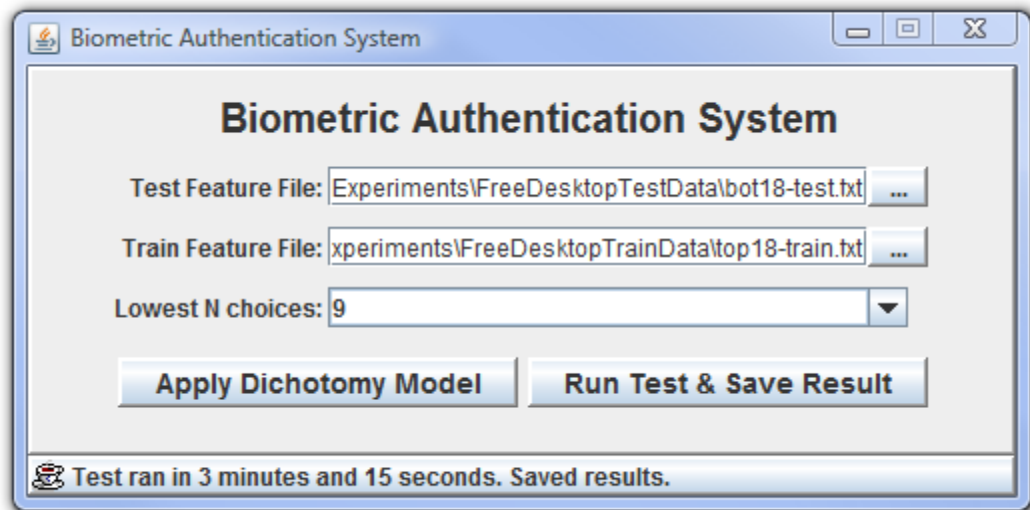


After a successful run, you should observe a screen similar to the following. Notice the time in the status bar and the 'Run Test & Save Result' button is now active.



Select the results directory and output file name.

- For Lowest N choices of 9 use:
c:\it691\keystroke A Experiments\FreeDesktopResults\freedesktop-9N.csv
- After the run is complete, you should see a screen similar to:

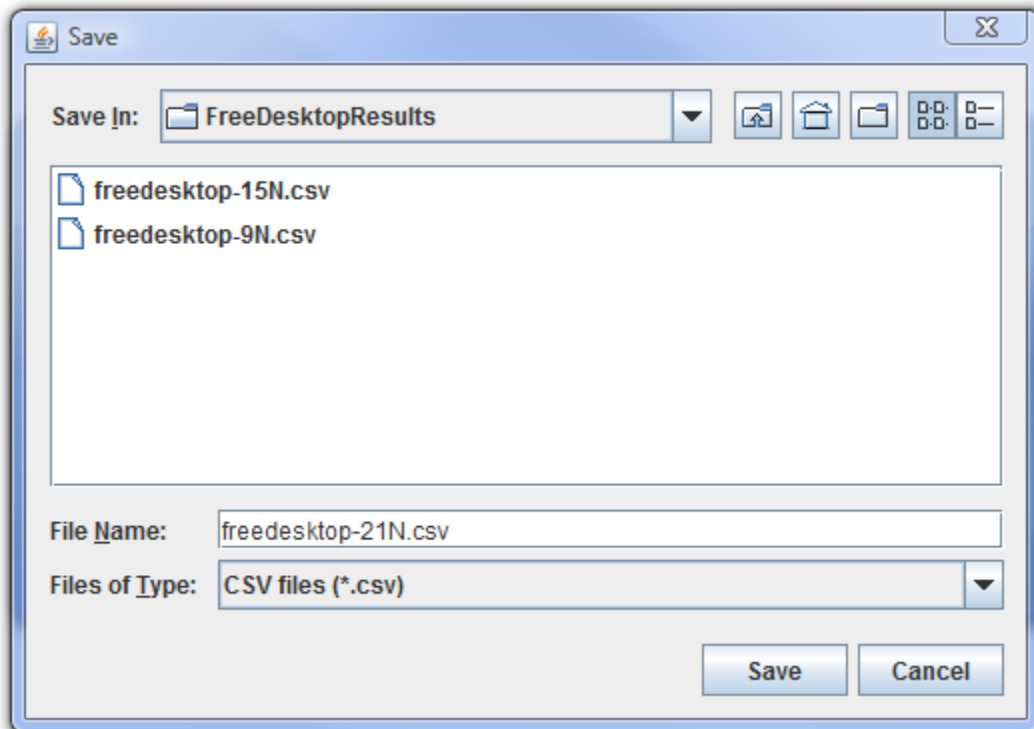
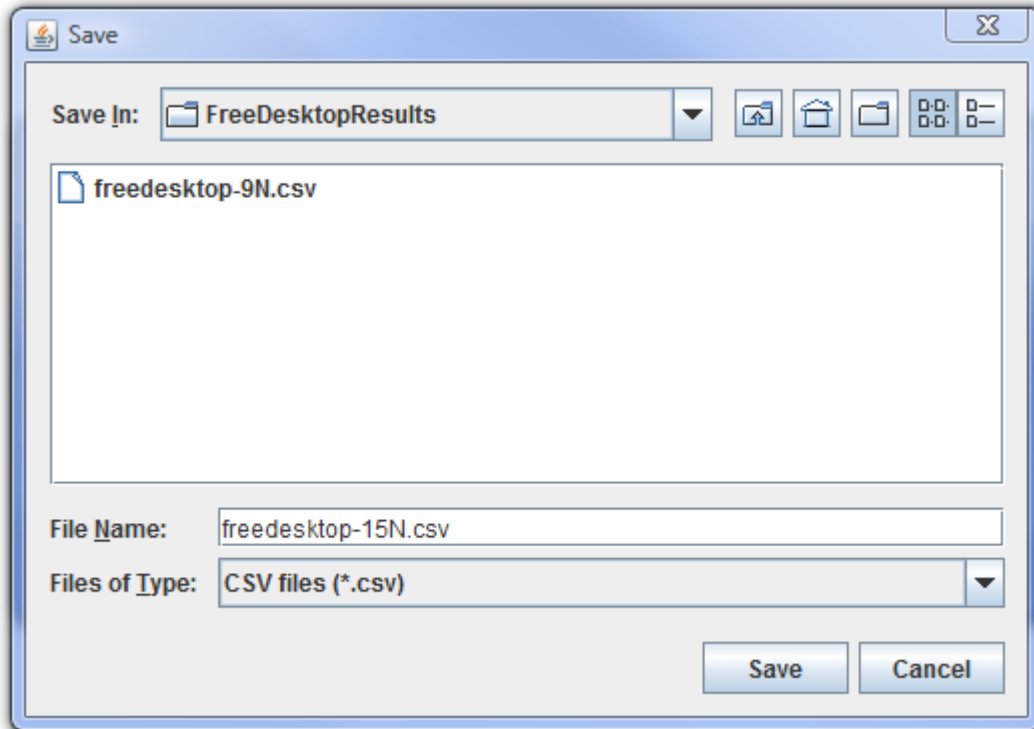


- c) Repeat above steps of 'Applying Dichotomy Model' and 'Running tests' using Lowest N choices of 15 and 21. Be sure to use the maximum amount of dichotomy data and use file names of:

C:\it691\keystroke A Experiments\FreeDesktopResults\freedesktop-15N.csv

C:\it691\keystroke A Experiments\FreeDesktopResults\freedesktop-21N.csv

- Applying the Dichotomy Model after the first time will result in very fast processing (0 seconds) for additional runs. This is expected behavior unless you change how much data should be dichotomized or delete the saved dichotomy data files.
- Be sure to select the desired output directory for 15N and 21N results and be careful not to overwrite existing data files. The default path is the location of the training file directory. Navigate to your results storage location:



- Laptop Free Experiments

Repeat all parts of the previous step using the LaptFree test and train data files created in the previous step.

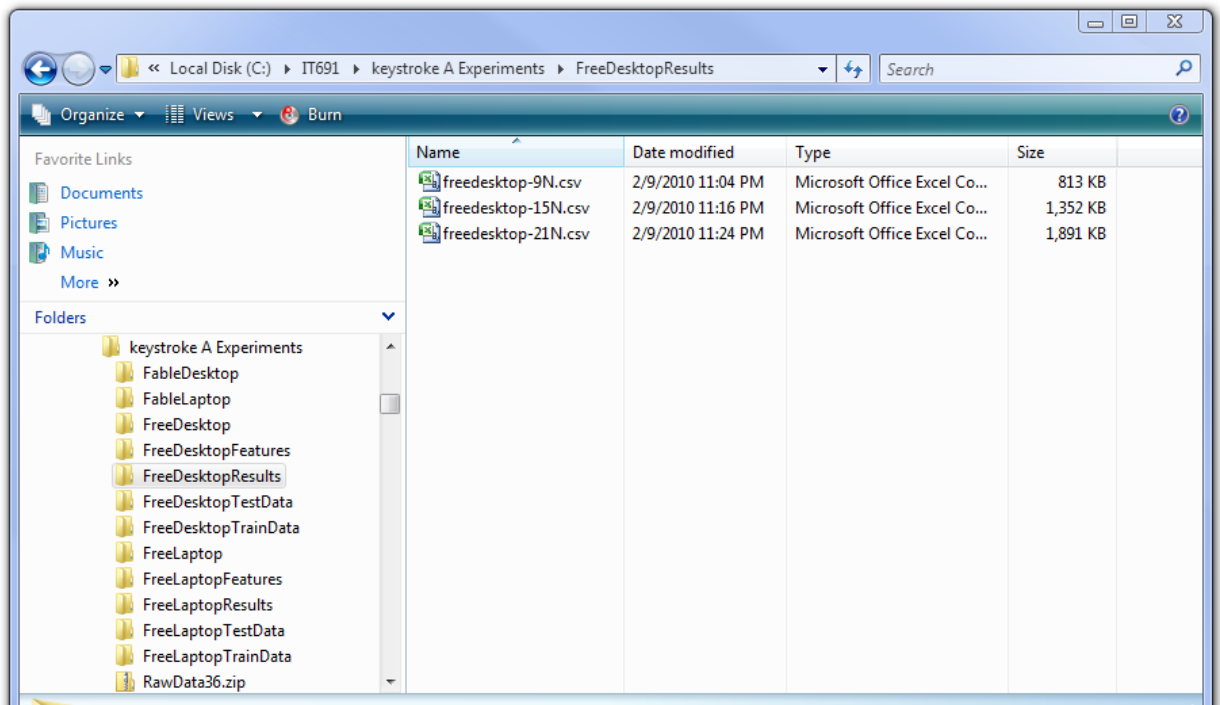
Run experiments using the 'Lowest N choices' of 9, 15, and 21.

A suggested output file directory is C:\it691\keystroke A Experiments\FreeLaptopResults

- Next-Steps

The obtained results will be used in the next deliverable, KBS Experiments 'B'. Before proceeding, please be sure you have created the following files successfully for desktop and laptop experiments:

(Laptop shown below)



The 9N file will contain results with data blocks containing 9 rows each, similar to:

```

freedesktop-9N.csv
-----
Keystroke,DESKTOP/FREETEXT7 1,165-3576,DESKTOP/FREETEXT1 1,176-3740,19,4,58,18,4,94
.1/W.
1,W,1,0606371074626428
2,W,1,1530331309265809
3,W,1,1543830783470175
4,W,1,2603603260750202
5,W,1,2752502505385437
6,W,1,3639707999451307
7,W,1,3985179916243569
8,W,1,4062045078134315
9,W,1,438417108047038
.2/W.
1,W,1,2164400848312176
2,W,1,218998999535185
3,W,1,3238655164376805
4,W,1,3327346677829024
5,W,1,372261730262886
6,W,1,4005054490887898
7,W,1,4513778721686372
8,W,1,4921851660916068
9,W,1,5136893785303822
.3/W.
1,W,0,9965091441149647
2,W,1,1373335132335425
3,W,1,1668783945210817
4,W,1,1908853462115614
5,W,1,2755008233413367
6,W,1,3342042191749515
7,W,1,3486728163630246
8,W,1,3611102701997768
9,W,1,3895815943439296
.4/W.

```

The N15 file will contain results with data blocks containing 15 rows each, similar to:

```

freedesktop-15N.csv
-----
Keystroke,DESKTOP/FREETEXT7 1,165-3576,DESKTOP/FREETEXT1 1,176-3740,19,4,58,18,4,94
.1/W.
1,W,1,0606371074626428
2,W,1,1530331309265809
3,W,1,1543830783470175
4,W,1,2603603260750202
5,W,1,2752502505385437
6,W,1,3639707999451307
7,W,1,3985179916243569
8,W,1,4062045078134315
9,W,1,438417108047038
10,W,1,4484697962413882
11,W,1,4632385857378758
12,W,1,469461931829818
13,W,1,4734108711744418
14,W,1,4761613284963482
15,W,1,480228746501633
.2/W.
1,W,1,2164400848312176
2,W,1,218998999535185
3,W,1,3238655164376805
4,W,1,3327346677829024
5,W,1,372261730262886
6,W,1,4005054490887898
7,W,1,4513778721686372
8,W,1,4921851660916068
9,W,1,5136893785303822
10,W,1,5632998709027488
11,W,1,5682647180516822
12,W,1,577812086332441
13,W,1,5812091216225954
14,W,1,589216088208049
15,W,1,5938143846762864
.3/W.
1,W,0,9965091441149647
2,W,1,1373335132335425
3,W,1,1668783945210817
4,W,1,1908853462115614
5,W,1,2755008233413367
6,W,1,3342042191749515

```

The N21 file will contain results with data blocks containing 21 rows each.

- Desktop results will be similar in structure.