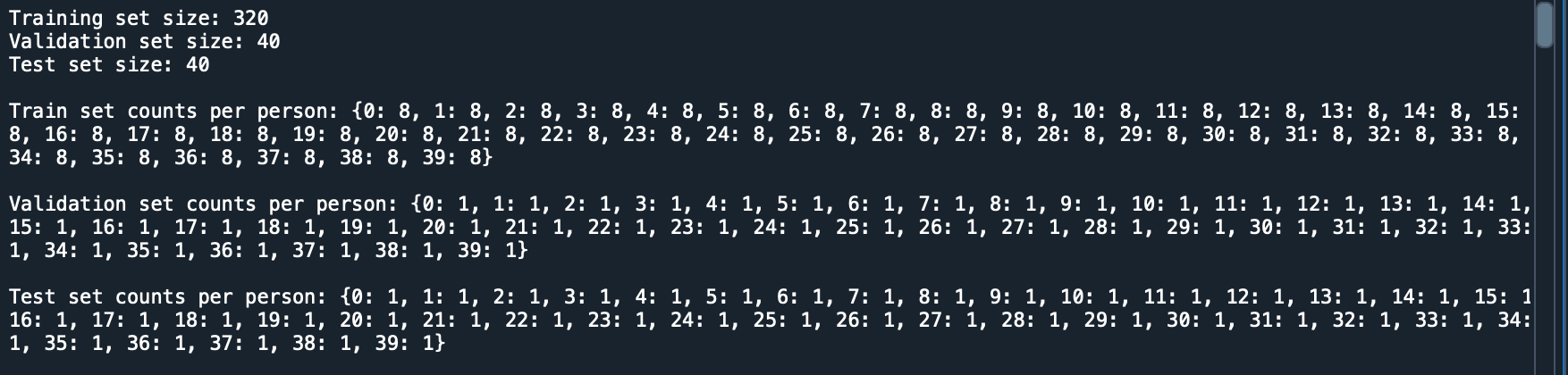
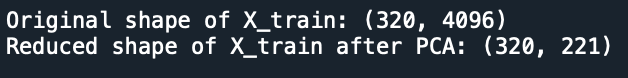
**Assignment 4 Written Report**

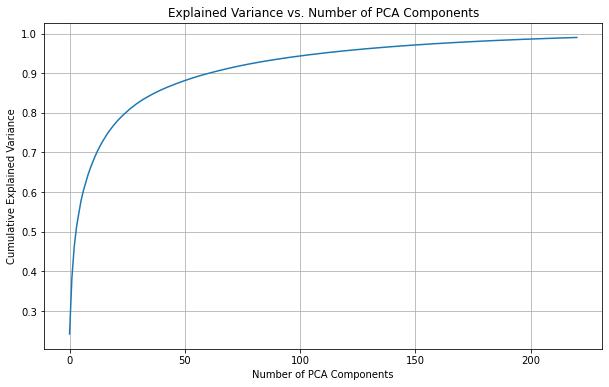
Student Name: Artem Kamov  
Student ID: 301220613

2. The split ratio is 80% for the training test and 20% for the testing and validation set. The split ratio for the testing and validation set is 0.5. This ratio was chosen because it provides a greater amount of data for the training set than other rations, which is useful in the given case, as there is limited data in the dataset (400 pictures). Besides, Stratified sampling was applied to ensure an equal number of images per person in each set.

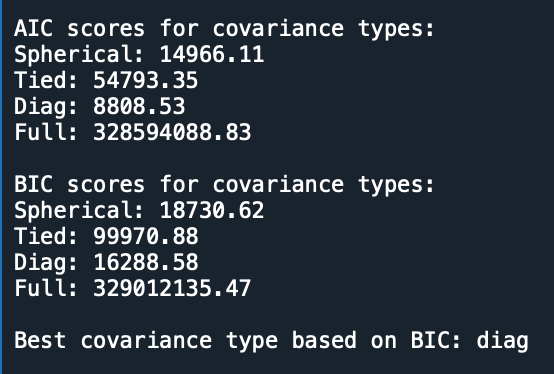


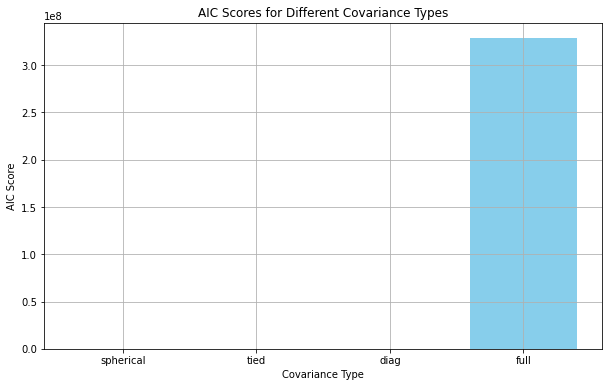
3. PCA is applied to the training data, preserving 99% of the variance, to reduce the dataset’s dimensionality.

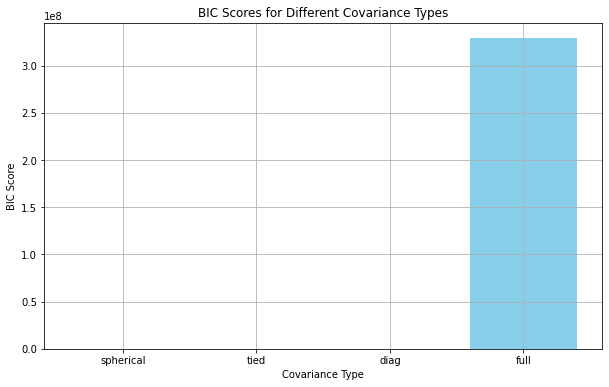




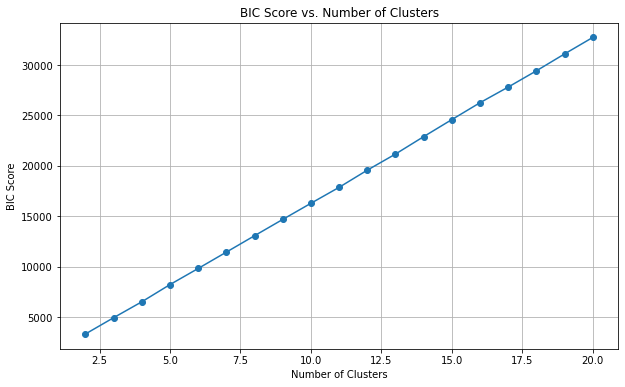
4. As we can see, the the most suitable covariance type for the dataset is diagonal, since it has the smallest AIC and BIC scores of 16288.58. We also can observe a great difference between AIC and BIC scores for each covariance which is likely due to the complexity of the models, especially with full and tied covariance types. They require many more parameters to describe the dataset. In contrast, simpler covariance types (like spherical and diag) require fewer parameters and therefore yield significantly smaller AIC/BIC values.



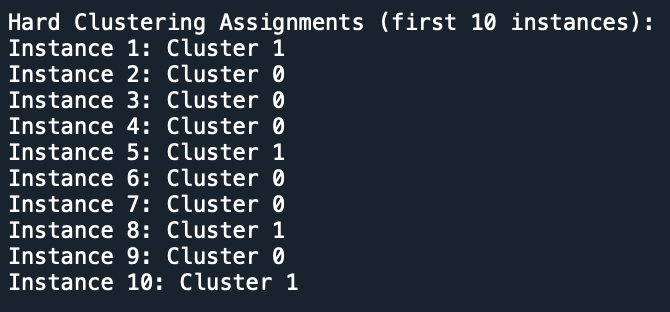


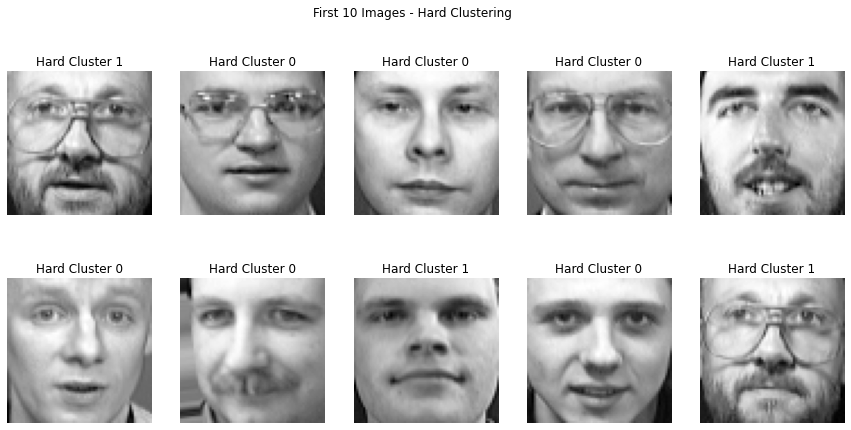


5. BIC was used to determine the minimum number of clusters that best represent the dataset. According to the BIC, the optimal number of clusters is 2.

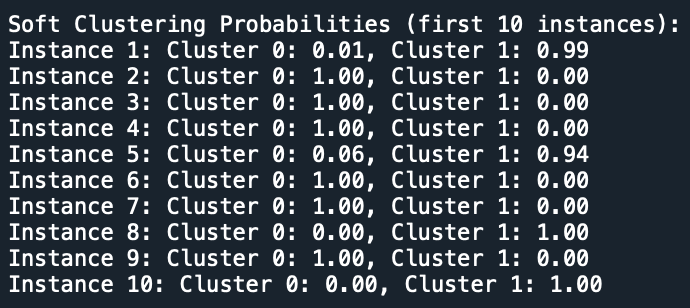
****

7.





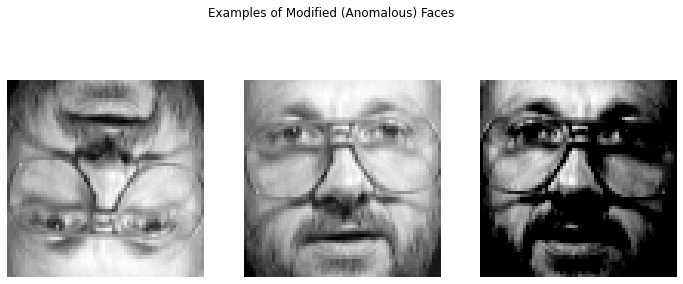
8.





9. Generation of 20 new face images using the sample() method:  


10. 20 images were modified. Now, each of these images has 3 modified versions - 180 degree rotated, horizontally flipped, and darkened. Example of a modified image:



11. Based on the log likelihood values for both normal and modified images, it appears that the model could detect many of the anomalies, since the score\_samples() method indicates some differences in log likelihoods between normal images and modified (anomalous) versions

So, the log likelihoods for rotated images are significantly lower than those for the normal images. For example, Image 1's normal score is -27.80, whereas the rotated version has a score of -225.16. This large drop in score indicates that the rotated images deviate strongly from the patterns learned by the model making them identifiable as anomalies.

In contrast, the flipped images vary in their likelihoods. Some flipped images, such as Image 5 (-38.58) and Image 6 (11.71), have unusually similar likelihoods compared to their normal counterparts (-38.27 and -10.66 respectively). This suggests that flipping the images doesn’t always produce a noticeable anomaly for the model, likely because the flipped version is still somewhat similar to the normal pattern.

Meanwhile, darkened images consistently show lower log likelihood scores than normal images, though the drops are not always as drastic as with rotated images. For instance, Image 1 darkened has a likelihood of -45.49, which is noticeably lower than its normal likelihood of -27.80, suggesting that the model can partially recognize this as an anomaly but is not as sensitive to darkening as it is to rotation.

Overall, the model appears reasonably effective at identifying certain types of anomalies, especially rotations and darkening. Flipping seems to be less effective in triggering low likelihood scores which indicates that this modification may not be as easily detectable by the model due to its similarity to the normal image pattern.

Log Likelihoods for Normal Images:

Image 1: -27.80

Image 2: -12.60

Image 3: 6.38

Image 4: -19.54

Image 5: -38.27

Image 6: -10.66

Image 7: -2.95

Image 8: -29.29

Image 9: -14.08

Image 10: -52.94

Image 11: -51.85

Image 12: -43.04

Image 13: 25.56

Image 14: 6.76

Image 15: -54.12

Image 16: 34.71

Image 17: -53.36

Image 18: -44.80

Image 19: -56.70

Image 20: -53.42

Log Likelihoods for Modified (Anomalous) Images:

Image 1 rotated: -225.16

Image 1 flipped: -32.05

Image 1 darkened: -45.49

Image 2 rotated: -120.83

Image 2 flipped: 48.77

Image 2 darkened: -70.05

Image 3 rotated: -167.86

Image 3 flipped: 18.42

Image 3 darkened: -35.22

Image 4 rotated: -160.94

Image 4 flipped: 6.42

Image 4 darkened: -39.76

Image 5 rotated: -293.18

Image 5 flipped: -38.58

Image 5 darkened: -44.19

Image 6 rotated: -90.15

Image 6 flipped: 11.71

Image 6 darkened: -31.66

Image 7 rotated: -139.83

Image 7 flipped: -8.20

Image 7 darkened: -24.54

Image 8 rotated: -228.77

Image 8 flipped: -12.34

Image 8 darkened: -59.16

Image 9 rotated: -299.54

Image 9 flipped: -10.83

Image 9 darkened: -14.29

Image 10 rotated: -161.00

Image 10 flipped: -19.20

Image 10 darkened: -44.64

Image 11 rotated: -100.81

Image 11 flipped: -26.76

Image 11 darkened: -69.76

Image 12 rotated: -376.02

Image 12 flipped: -39.54

Image 12 darkened: -49.67

Image 13 rotated: -135.62

Image 13 flipped: 22.94

Image 13 darkened: -22.91

Image 14 rotated: -146.87

Image 14 flipped: 5.06

Image 14 darkened: -27.92

Image 15 rotated: -154.81

Image 15 flipped: -65.68

Image 15 darkened: -73.92

Image 16 rotated: -99.14

Image 16 flipped: 46.85

Image 16 darkened: -26.21

Image 17 rotated: -159.89

Image 17 flipped: -30.90

Image 17 darkened: -54.49

Image 18 rotated: -130.04

Image 18 flipped: -36.24

Image 18 darkened: -47.35

Image 19 rotated: -134.04

Image 19 flipped: -17.87

Image 19 darkened: -54.20

Image 20 rotated: -121.82

Image 20 flipped: -68.45

Image 20 darkened: -30.45