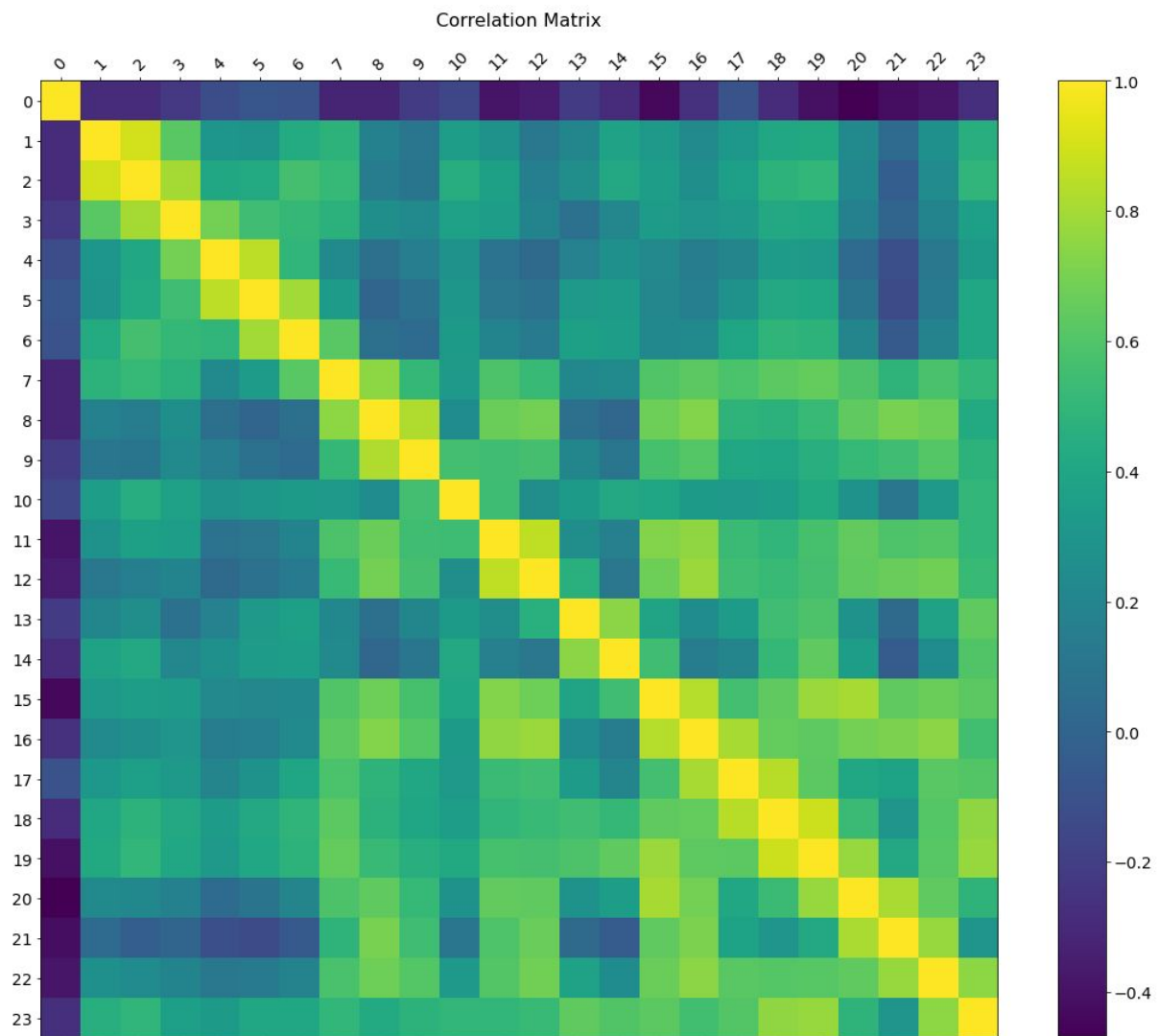


INF2B Learning CW Report Task 1

Linda Mazánová
S1801828

1.2 Describing the correlation matrix

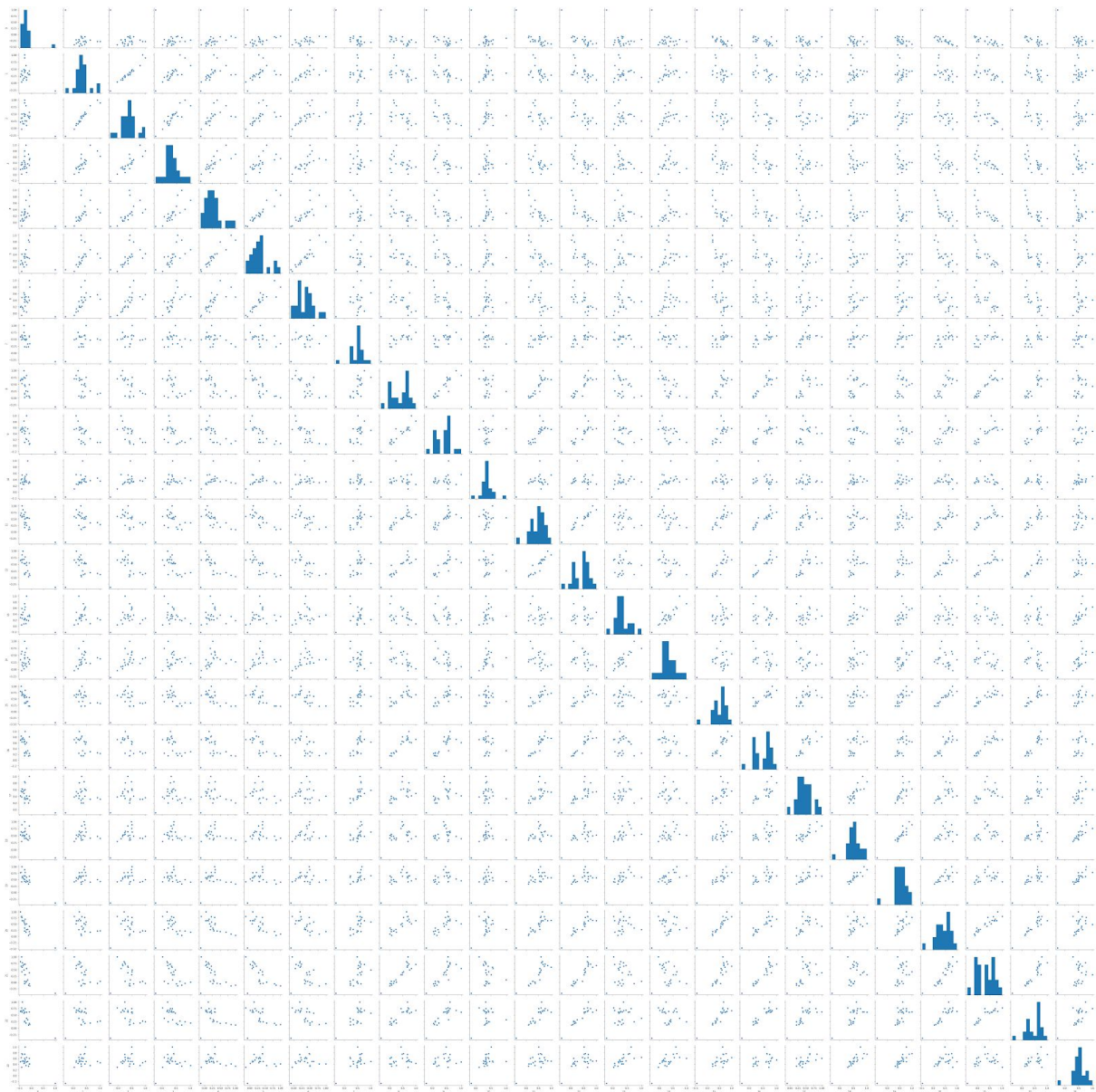


The dimensions of the correlation matrix are 24x24 since the dimension of all feature vectors is $D=24$. We can observe certain patterns:

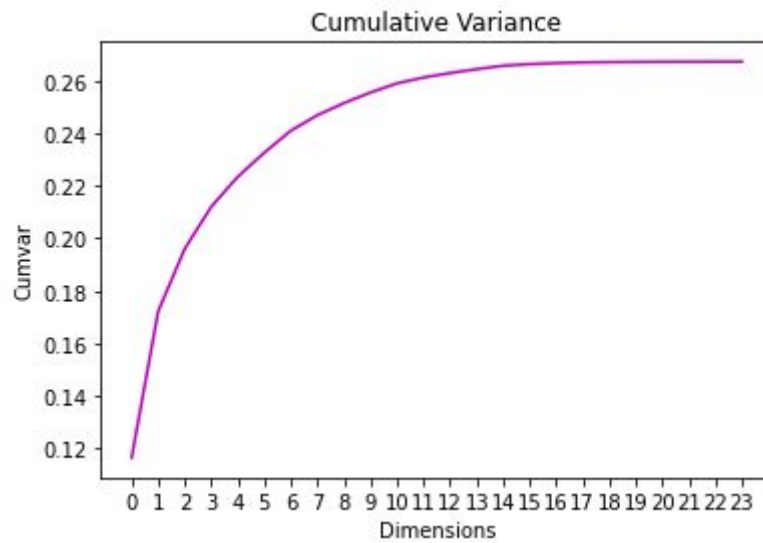
1. Most of the dimensions that are 1 apart, e.g. $D=10$ and $D=9$ show positive correlation. Not all adjacent dimensions show the same strength of correlation, for example $D=2$ and $D=1$ are more strongly positively correlated (shown by more yellow color) than $D=10$ with $D=11$ (shown by darker green color).
2. There is a yellow line $y=-x$ that shows positive correlation of value 1.0 since any dimension positively correlates with itself.

3. Starting from the line $y=-x$, the further away from each other the dimensions are, the weaker the positive correlation. It changes to negative correlation (e.g. $D=21$ and $D=2$).
4. When $x=0$ or $y=0$ we see negative correlation, for example $D=0$ with $D=22$. Note that the negative correlation is not as strong, the absolute value of this negative correlation never reaches 0.5.

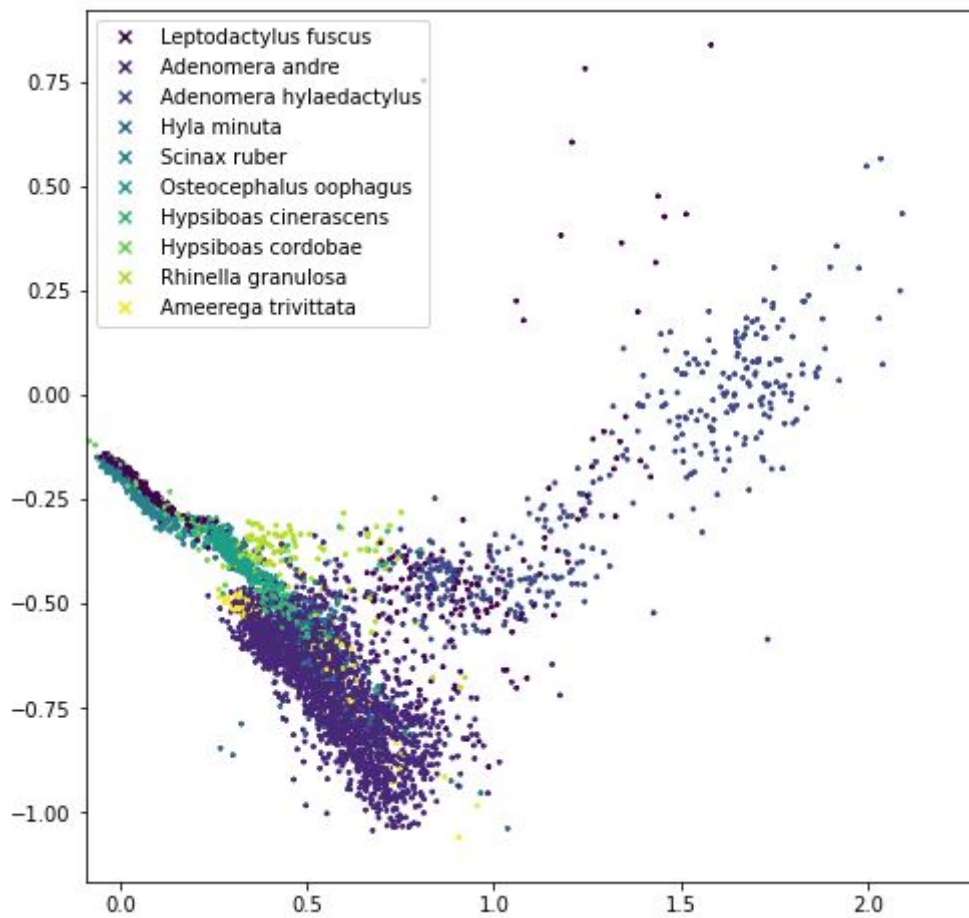
We can also plot and observe how each dimension correlates with all the other dimensions by looking at the plot that plots pairs of dimensions. We observe that on the left and top edge there tends to be negative correlation of features so all other dimensions negatively correlate with $D=0$. As we go closer to the line $y=-x$, the correlation becomes positive and more strong.



1.3b Graph of Cumulative Variance with respect to the dimensions



1.3c Data of different classes shown on a 2D PCA plane



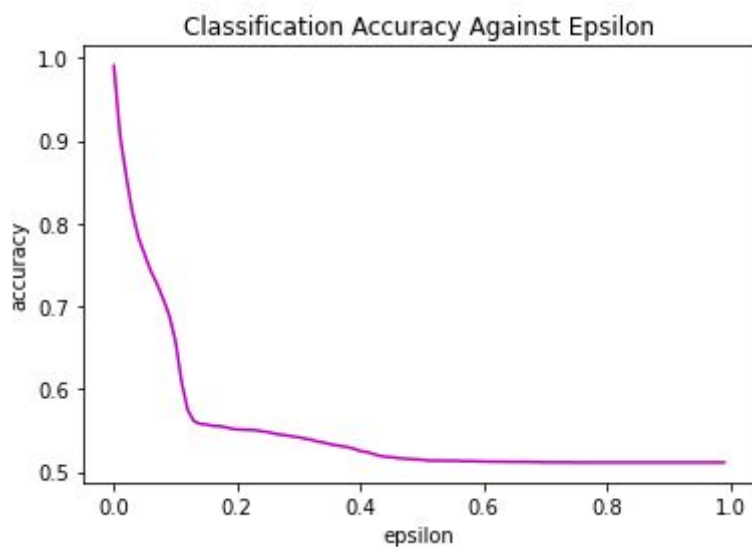
1.4 Correct classification rates for 3 types of covariance matrix:

Accuracy for CovKind=1:
0.9065539578937036

Accuracy for CovKind=2:
0.8278328507631152

Accuracy for CovKind=3:
0.8807716215995157

1.5 A graph that shows the classification accuracy with respect to the regularisation parameter epsilon:



As the value of epsilon increases by a very small amount the accuracy drops significantly. We see a rapid drop in accuracy as we increase epsilon from a very small number, close to zero, up to approximately 0.12. After this value of epsilon, further increase in the epsilon parameter still lowers the accuracy, but the rate of lowering the accuracy is much slower. After the value of epsilon reaches approximately 0.45 the accuracy remains fairly constant and further increase in the epsilon parameter does not change the accuracy. The graph eventually reaches a plateau and starts behaving asymptotically.