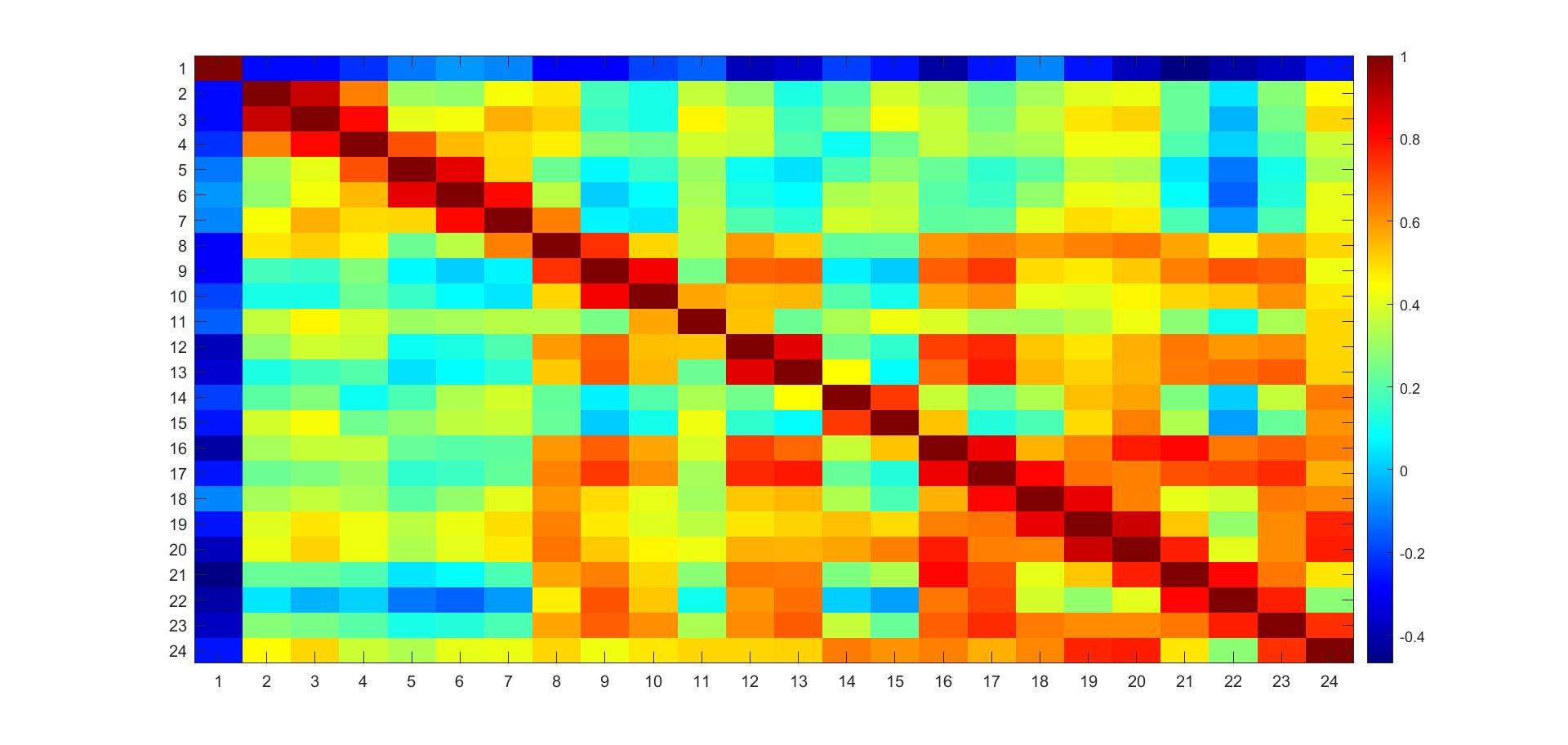
**Name: Liu Rongxing**

**UUN: s1810054**

**Informatics 2b——Learning Coursework 1 Task1**

**Task1\_2:**

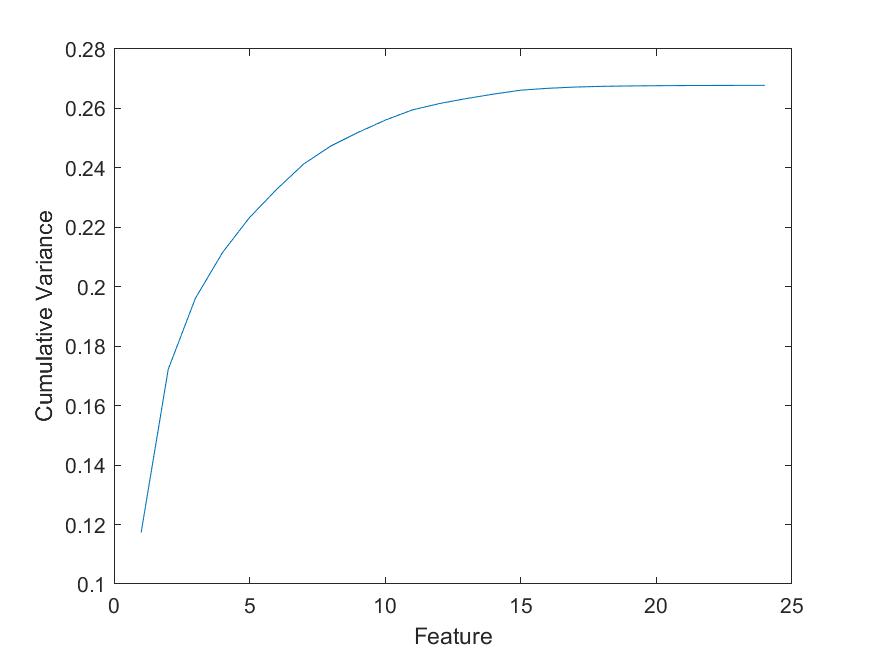
The correlation matrix can be presented more visually with a colored graph.



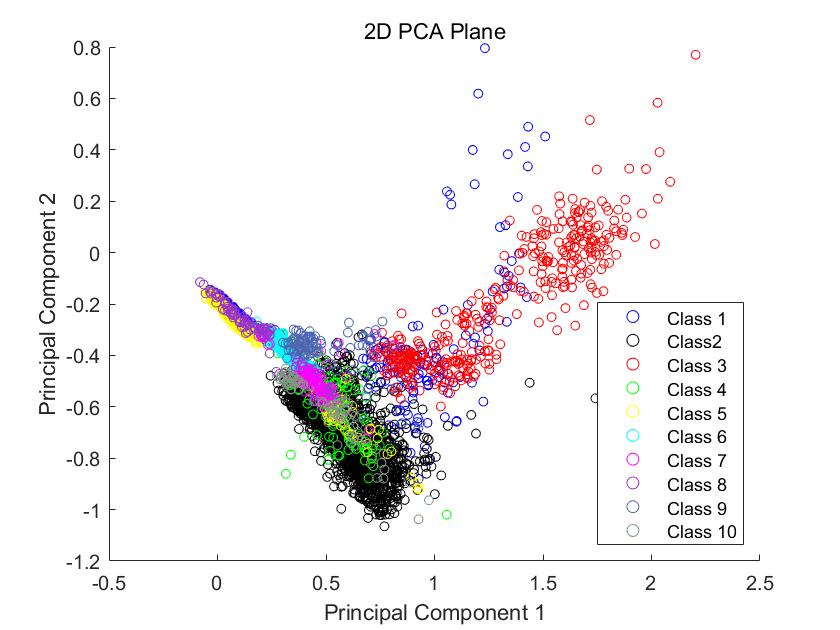
We can see from the graph that almost all the ‘neighboring’ features have rather strong correlations (with correlation coefficient being close to or exceeding 0.8) except for **Features (13,14), (15,16), (7,8), (1,2), (10,11) and (11,12).** In addition, **Features (24,19), (24,20), (16,20), (16,21) and (13,17)** also have strong positive correlations.

As for the negative correlation, we can see that **Feature 1** is almost negatively correlated to all other features. However, the all the negative correlations are not strong because they are all above -0.5.

**Task1\_3: Graph of Cumulative Variance**



**Task1\_3: Data on 2D-PCA**

****

**Task1\_4**

When run with **Covkind=1,** the accuracy is 90.82%.

When run with **Covkind=2,** the accuracy is 79.55%.

When run with **Covkind=3,** the accuracy is 90.39%.

**Task1\_5**

From the graph, we can see that **as the value of epsilon increases, the accuracy decreases**.

However, as the accuracy approaches 0.74, it will stay around this value.

