Task3 – SVM + ANOVA

```
→  task3 python3 main.py
(6500, 14)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1941
accuracy_count:  1654
accuracy_score:  0.8521380731581659
precision score:  0.8516209612044143
recall score:  0.8506809858728822
執行時間: 24.381994009017944 秒
```

```
→  task3 python3 main.py
(6500, 7)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1933
accuracy_count:  1629
accuracy_score:  0.8427315054319711
precision score:  0.842057040330853
recall score:  0.8416219644263194
執行時間: 24.273701667785645 秒
```

```
→  task3 python3 main.py
(6500, 6)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1935
accuracy_count:  1571
accuracy_score:  0.8118863049095607
precision score:  0.8137745001741458
recall score:  0.8108958755451338
執行時間: 21.966783046722412 秒
```

We use **Top 7 features**, and [1, 7, 8 ,9, 10, 11, 13, 14] are not important in our SVM Model

```
→  task3 python3 main.py
[ 0  2  3  4  5  6 12]
[3.32917803e+00 6.60045825e-01 1.07905271e+03 1.04430877e+03
 7.11083325e+02 7.24830073e+02 3.69187489e+02 2.14698042e+00
 7.83977089e-01 1.99358098e+00 4.48417825e-01 1.00328539e-01
 2.18087360e+00 2.17806063e-01]
```

ANOVA 判斷線性關係，
但若是資料不是線性的就會誤判

Task4 – SVM + ANOVA

```
→  task4 python3 main.py
(6500, 14)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1909
accuracy_count:  1369
accuracy_score:  0.7171293871136721
precision score:  0.7138799024293245
recall score:  0.7130493976808411
執行時間: 79.66132235527039 秒
```

```
→  task4 python3 main.py
(6500, 11)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1908
accuracy_count:  1319
accuracy_score:  0.6912997903563941
precision score:  0.6903801471725978
recall score:  0.687036559521707
執行時間: 81.34038949012756 秒
```

```
→  task4 python3 main.py
(6500, 10)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1885
accuracy_count:  1247
accuracy_score:  0.6615384615384615
precision score:  0.65860787560396
recall score:  0.6568400470690146
執行時間: 79.33542513847351 秒
```

We use Top 11 features, and [0, 9, 12] are not important in our SVM Model

```
[ 1  2  3  4  5  6  7  8 10 11 13]
[1.68471472e-01 3.75465662e-01 4.14061149e+02 4.24059808e+02
 7.41531569e-01 5.52775655e+00 9.37109841e-01 7.84529139e-01
 6.16612839e-01 2.70993638e-01 9.25383184e-01 2.28952376e+00
 2.99013156e-01 1.08609632e+00]
(6500, 11)
```

Task3 – SVM + Mutual Information

```
→  task3 python3 main.py
(6500, 14)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1941
accuracy_count:  1654
accuracy_score:  0.8521380731581659
precision score:  0.8516209612044143
recall score:  0.8506809858728822
執行時間: 24.381994009017944 秒
```

```
→  task3 python3 main.py
(6500, 9)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1926
accuracy_count:  1594
accuracy_score:  0.8276220145379024
precision score:  0.8280925801245038
recall score:  0.8259831931164517
執行時間: 26.0490882396698 秒
```

```
→  task3 python3 main.py
(6500, 8)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1924
accuracy_count:  1575
accuracy_score:  0.8186070686070686
precision score:  0.8193633861872696
recall score:  0.8175503692040107
執行時間: 24.643512725830078 秒
```

We use **Top 10 features**, and [0, 8, 11, 12] are not important in our SVM Model

```
→  task3 python3 main.py
[ 1  2  3  4  5  6  7  9 10 13]
[0.00000000e+00 3.90512977e-03 2.16943118e-01 2.17183440e-01
 1.84900960e-01 1.63298894e-01 1.13850809e-01 1.94088476e-04
 0.00000000e+00 7.09843638e-03 9.06065965e-03 0.00000000e+00
 0.00000000e+00 0.00000000e+00]
(6500, 10)
```

Mutural information score是 nonparametric (非參數) 的，
因此能夠測量非線性資料之間的關係

Task4 – SVM + Mutual Information

```
→  task4 python3 main.py
(6500, 14)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1910
accuracy_count:  1388
accuracy_score:  0.7267015706806282
precision score:  0.7224080532235403
recall score:  0.7218708902808207
執行時間: 76.45102429389954 秒
```

```
(6500, 11)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1901
accuracy_count:  1343
accuracy_score:  0.7064702788006313
precision score:  0.7073021674235314
recall score:  0.7015408684482616
執行時間: 75.36816620826721 秒
```

```
→  task4 python3 main.py
(6500, 10)
================
SVM
Best parameters: {'C': 100, 'kernel': 'rbf'}
Test data count:  1903
accuracy_count:  1295
accuracy_score:  0.6805044666316342
precision score:  0.6826722522638252
recall score:  0.6758433978994872
執行時間: 75.6852490901947 秒
```

We use **Top 11 features**, and [1, 8, 9] are not important in our SVM Model

```
→  task4 python3 main.py
[ 0  2  3  4  5  6  7 10 11 12 13]
[0.00636359 0.         0.10783834 0.12174162 0.01145359 0.02996968
 0.02387777 0.00415483 0.         0.         0.00157616 0.01308285
 0.         0.00779174]
(6500, 11)
```