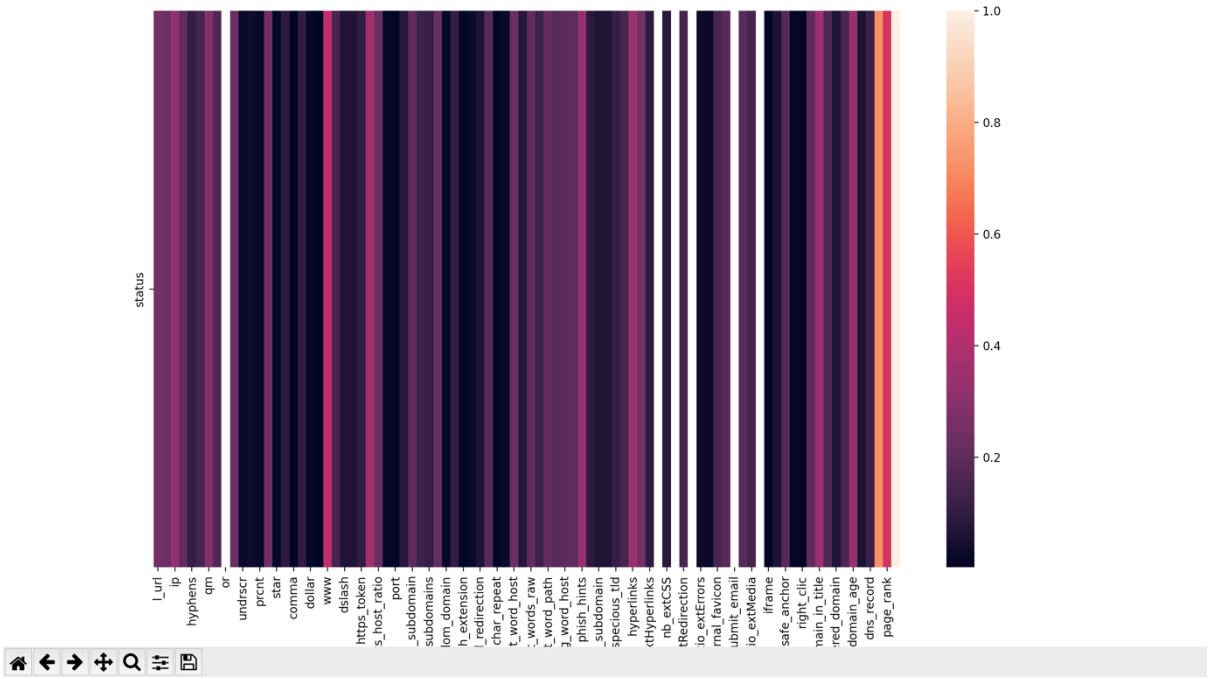
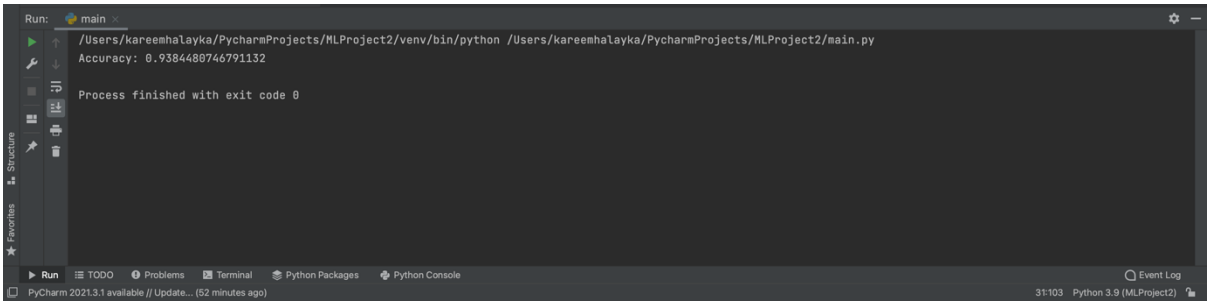
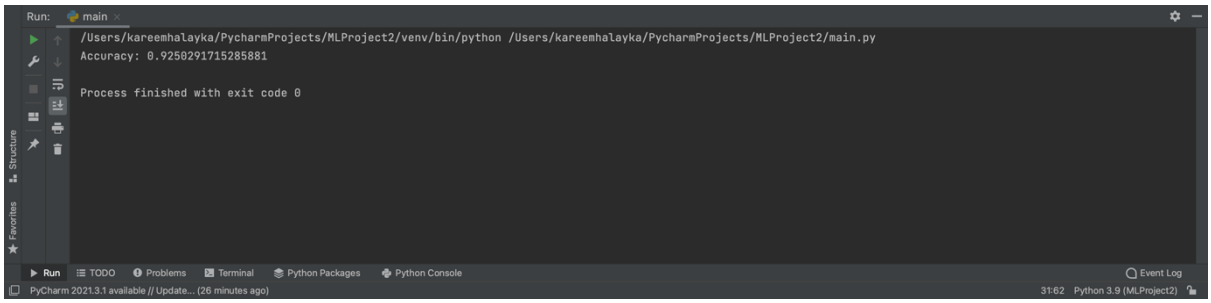
1- In this project, an SVM model was created to learn a dataset that includes 88 features. For the training phase, 8003 observations were used to train the SVM model. For the training phase, 3429 observations were used. In the beginning, with all features being used, the svm model was unusable, as it spent way too long running without giving any results. However, with proper tuning of features, the model gave great results in a considerably good time.
2- Firstly, a heatmap of the correlation between the status and all other features was created.

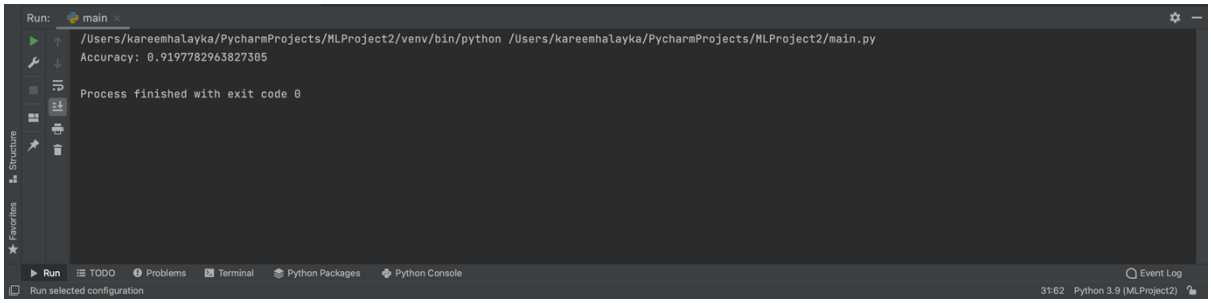


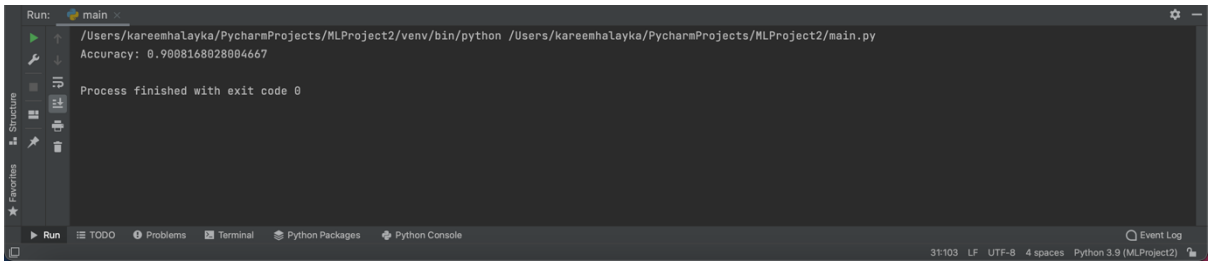
Secondly, a test was done to determine which features to be used to make SVM usable with good results. That was done by dropping the features with correlation near 0 in the beginning, then widening that range until the results became good. When features with correlation between 0.1 and -0.1 were dropped, the model took about 10 minutes to train with an accuracy of 93.8% for the test data.



Increasing that range to 0.2 and -0.2, 0.3 and -0.3, and 0.4 and -0.4 the accuracy dropped to near 90%, but the training time decreased to about 30 seconds on the same machine, which is a huge difference from the previous train time.







3- SVM is a popular classification algorithm that works on defining a boundary between the classes in a multidimensional space. It's often used for highly complex data as it can create boundaries in a high dimensional space that cannot be visualized. It can be split into two types: linear svm and none-linear svm. They both use a hyperplane (flat surface) to split the data, and this hyperplane's position is optimized using a margin that ensures that the hyperplane splits the data in the best way possible.

   The margin is defined by the distance between the datapoints at the two classes to be separated by the hyperplane. Optimization is done by finding the maximum margin, which is by locating the widest gap between the two closest points each from a class on a side of the hyperplane.

   Another way for determining the maximum margin is to search through space for all feasible hyperplanes in order to identify a pair of parallel planes that divide the points into homogenous groups and are as far apart as possible.

4- Results:

	precision	recall	f1-score	support
0	0.89	0.91	0.90	1715
1	0.91	0.89	0.90	1713
accuracy			0.90	3428
macro avg	0.90	0.90	0.90	3428
weighted avg	0.90	0.90	0.90	3428

bias:
variance:

from the values of bias and variance, we can conclude that in general there is no overfitting or underfitting in the data as the values are balanced.

Obada Tahayna 1191319
Karim Halayqa 1192087