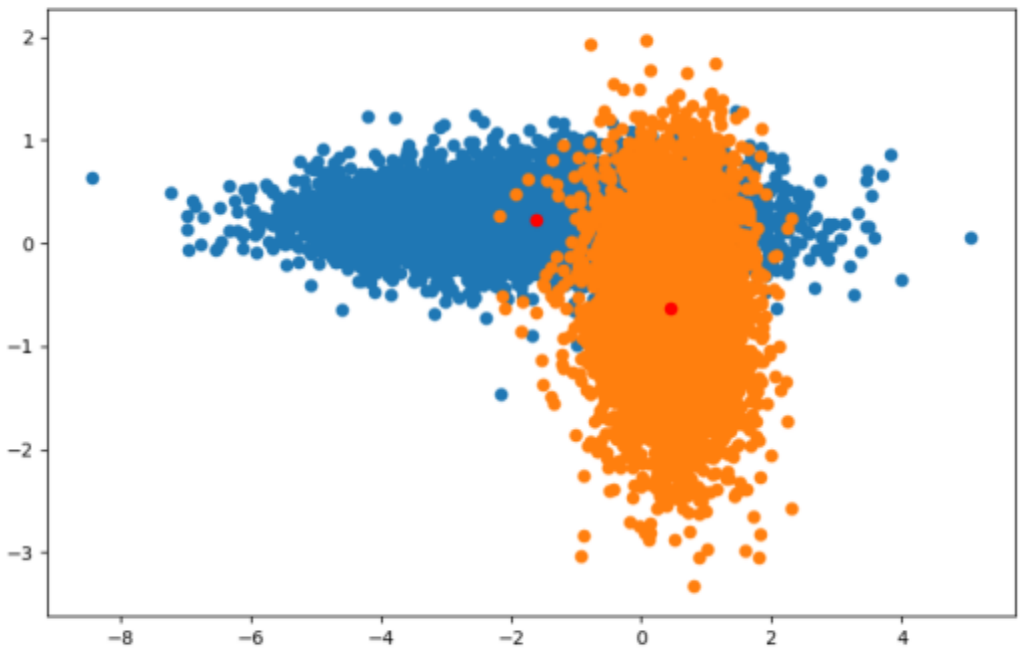
Lecture 3 group 743:

In task A we calculate the prior based on the train data. This gives us an accuracy of 90.82%

In task B we assume that either class are equally likely, and test this on a different dataset. Here we see an accuracy of 88.9%.

In task C we are given priors of the two classes, and using these get a higher accuracy of 96.11%

We discussed in the group what could cause the difference between B and C, and decided to check how many points were in each class, and found that there were significantly more points in class 1 than 2. There were 1135 points in class 1, and 126 in class 2. This fits with the 90% prior C1 and 10% prior C2



Plot of trainX(blue) and trainY(orange) with their medians shown with red dots