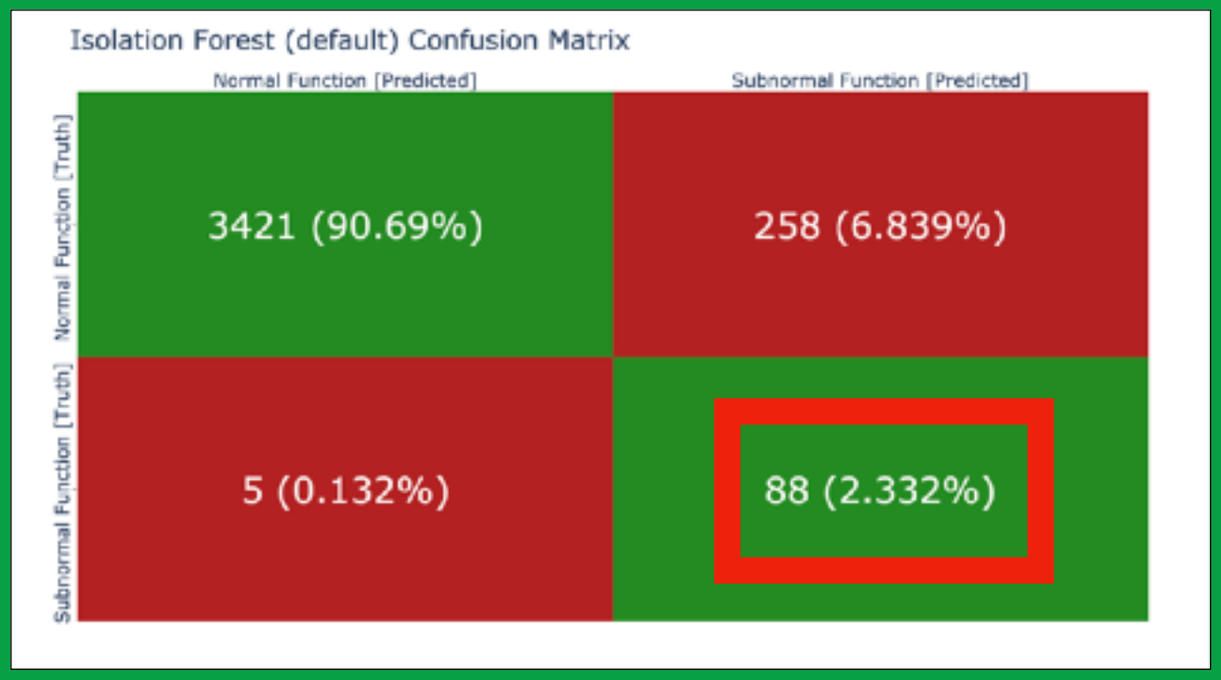
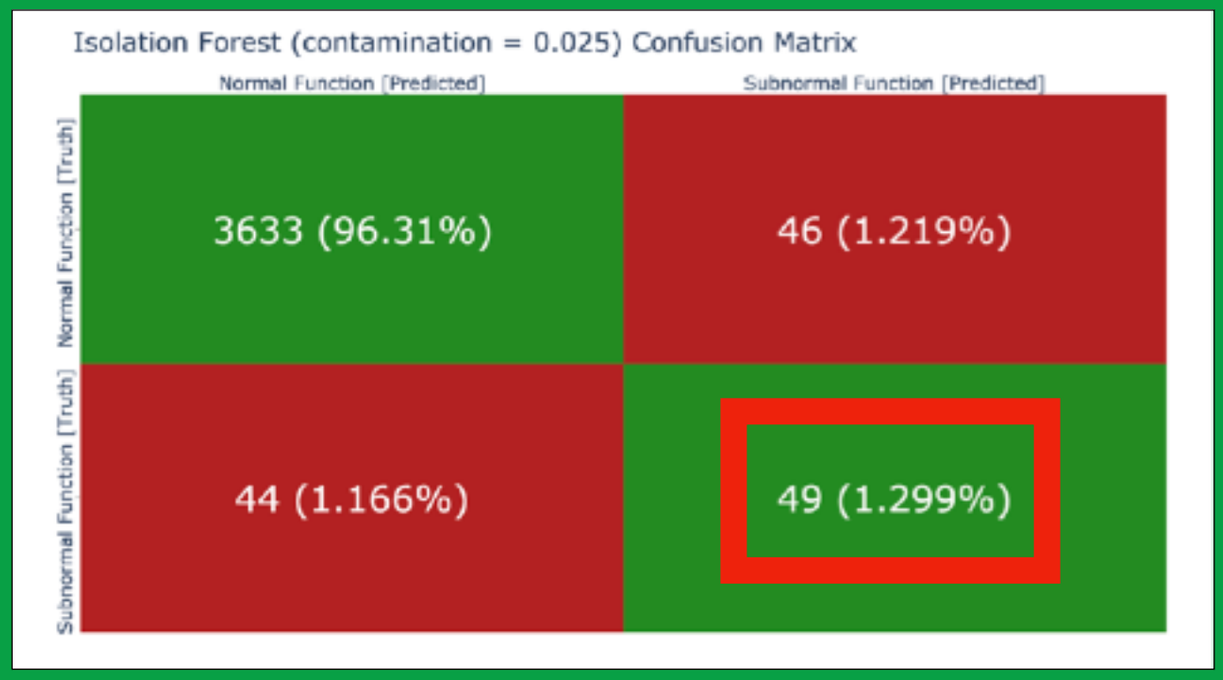
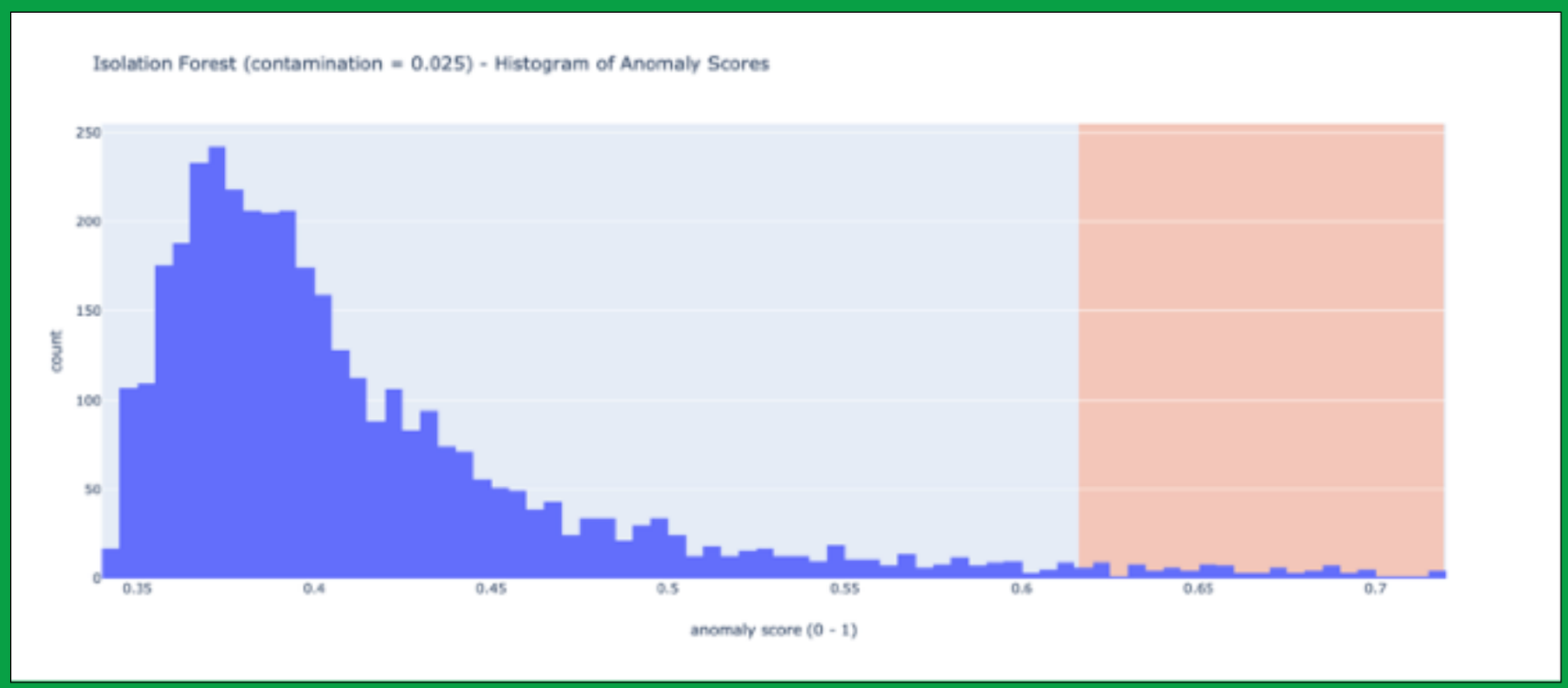
IF Performance







Overall, not bad…

Threshold gets automatically set to a questionable point by default, but as soon as you set the correct contamination value the results seem respectable.

About half of the outliers within the set have been identified correctly as outliers. The other half being misclassified as inliers, and a set of similar size becoming false positives just to make life more difficult - could be worse.

EIF Performance

Ummmmmm…. ?

The performance starts roughly the same when extension = 0, which is expected as this is just standard IF.

However, as the extension parameter increases…

i.e. we draw more complex, higher-dimension, hyperplanes

… we see a deterioration of outlier detection performance?

I'm going to do the noble thing and blame it on the data.