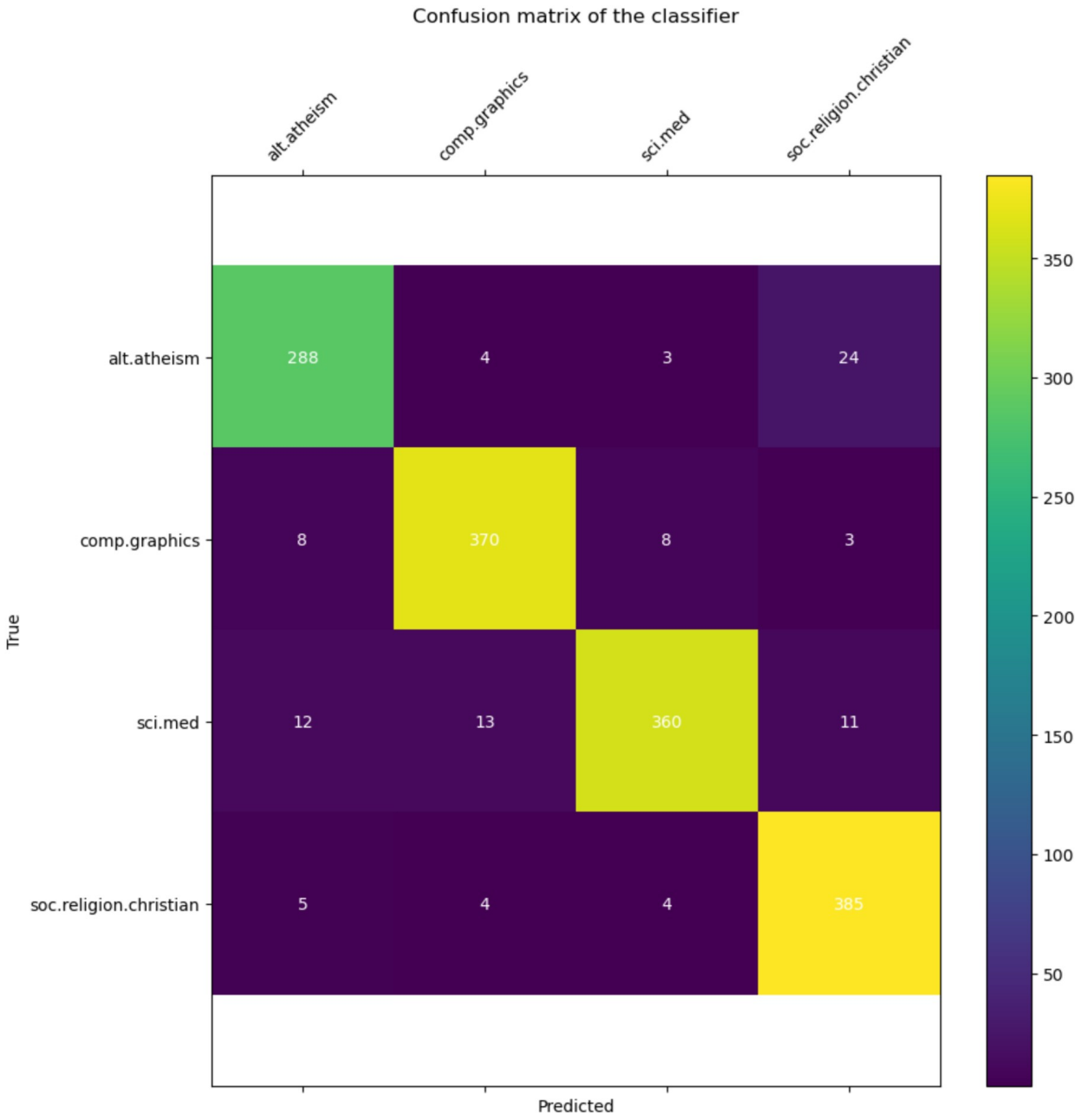
TP, TN, FP and FN

In our exercise in Unit 2, we have visualised the predictions of our model in a confusion matrix.

In the following we discuss how to calculate True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) values from this confusion matrix.

The first thing we need to bear in mind is that these four values are calculated from the perspective of each of the classes, i.e. in our problem with 4 classes, we will calculate each of these values 4 times.



Given the confusion matrix above, let's calculate the values for the alt.atheism class first. From the perspective of this class, atheism is the positive class whereas the remainder are negative classes:

- TP: these are the correct positive predictions (i.e. predicted alt.atheism, actually alt.atheism). Our TP for alt.atheism is 288 (diagonal value for this class).
- TN: these are the correct predictions across the negative classes. That is, all the values in the diagonal except for the alt.atheism class. That is, 370 + 360 + 385 = 1115.
- FP: these are the items that we predicted as positive, but were actually negative, i.e. anything that was predicted as alt.atheism, but should be any of the other three classes. These are the values in the column alt.atheism, except for the one in the diagonal which are the correct ones. That is, 8 + 12 + 5 = 25.
- FN: these are the items that we predicted as negative (either of the three other classes), but was actually positive (alt.atheism). These are the values in the row alt.atheism, except for the one in the diagonal. That is, 4 + 3 + 24 = 31.

Those are the 4 scores for alt.atheism. Now, let's move on to comp.graphics:

- TP: predicted and actual comp.graphics, in the diagonal, that's 370.
- TN: predicted and actual any of the other 3 classes, diagonal values except comp.graphics, that's 288 + 360 + 385 = 1033
- FP: predicted comp.graphics, but should have been something else, i.e. values in column comp.graphics, without the diagonal value. That's 4 + 13 + 4 = 21
- FN: predicted any of the other 3 classes, but it should have been comp.graphics, i.e. value in row comp.graphics, without the diagonal value. That's 8 + 8 + 3 = 19

And we would do the same for the other two classes.