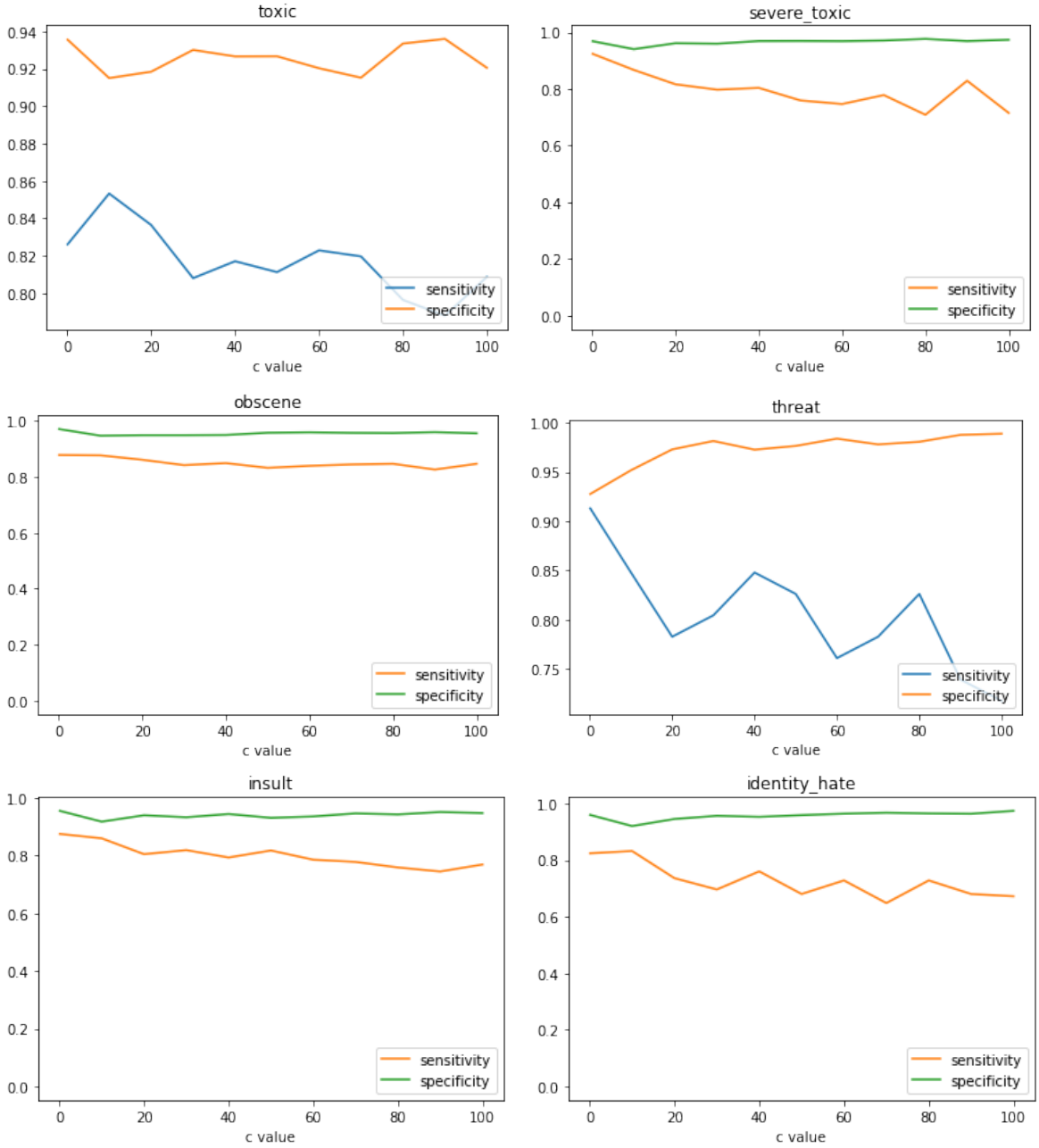
Sensitivity (true positive rate) = TP / (TP + FN)
Specificity (true negative rate) = TN / (TN + FP)

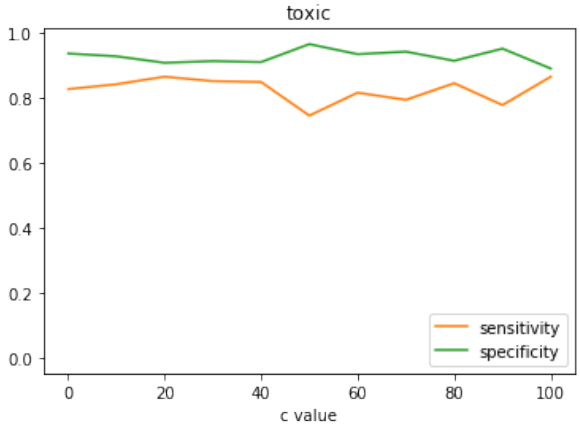
I assume our priority is sensitivity, is that correct? Since we don't want too many toxic comments to escape? In addition, specificity value is always quite high (>0.9), so I think the main issue is to increase sensitivity.

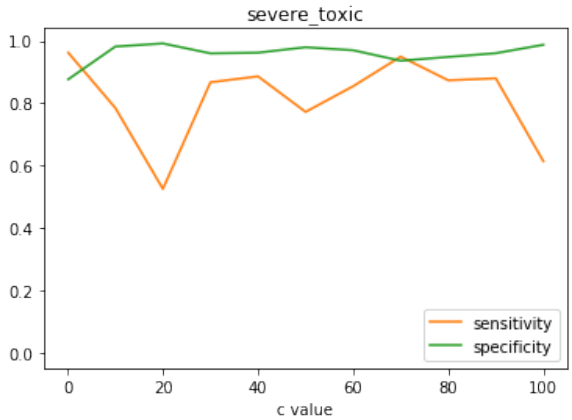
Setting class_weight to 'balanced' greatly increases sensitivity for all solvers (maximum sensitivity usually around 0.85-0.9).

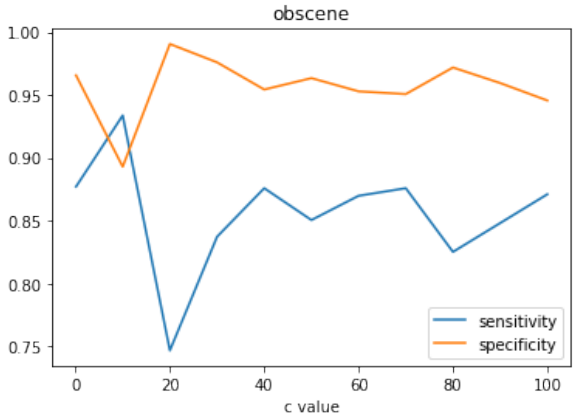
class_weight = 'balanced', solver='sag', other parameters set to default:

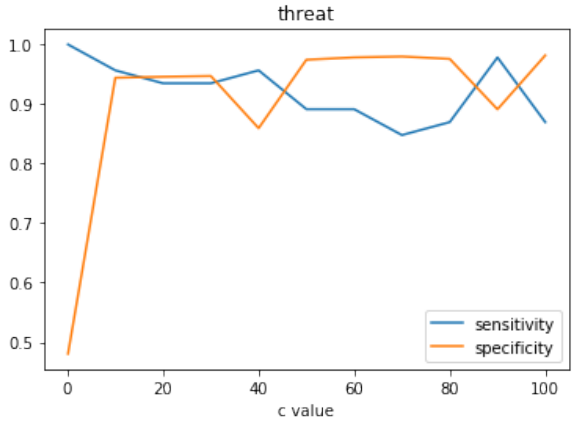


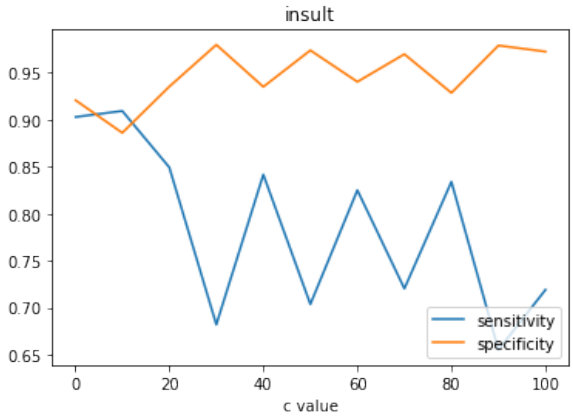
class_weight = 'balanced', solver='saga', other parameters set to default:

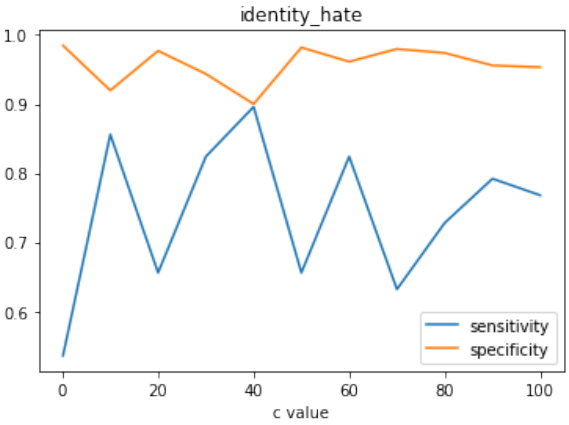






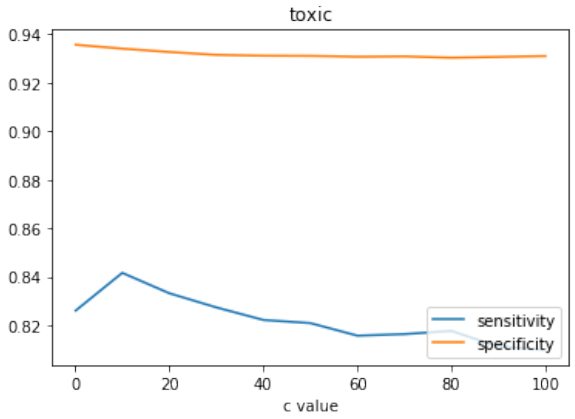


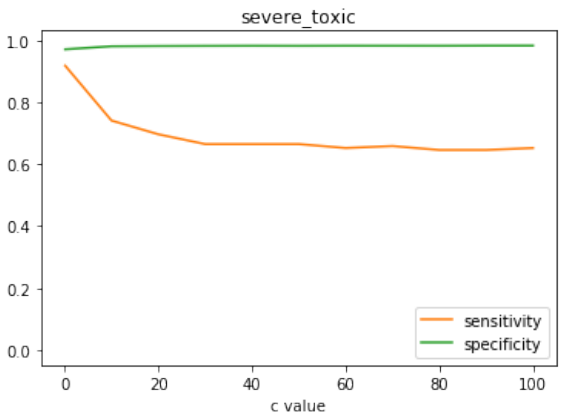


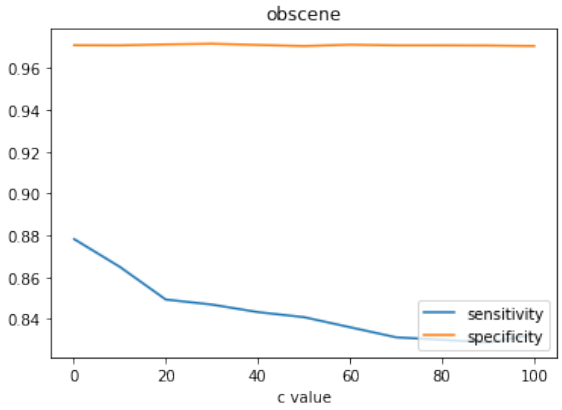


The two other types of solver ‘newton-cg’ and ‘lbfgs’ also produce similar (slightly worse) results. Didn’t finish drawing all the graphs, here are a few for comparison:

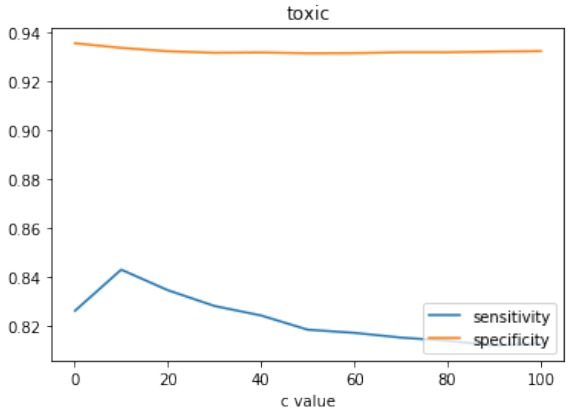
class_weight = ‘balanced’, solver=’lbfgs’, other parameters set to default:







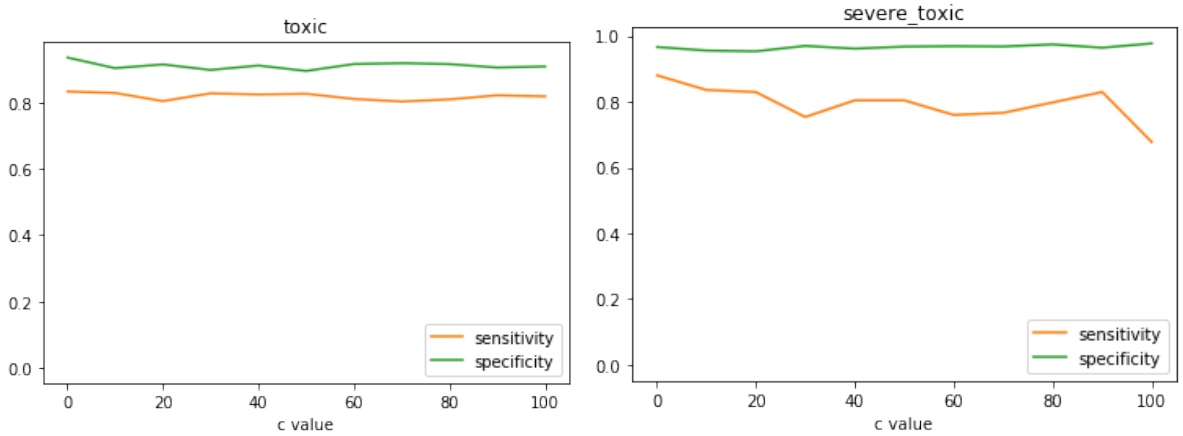
class_weight = ‘balanced’, solver= ‘newton-cg’, other parameters set to default:



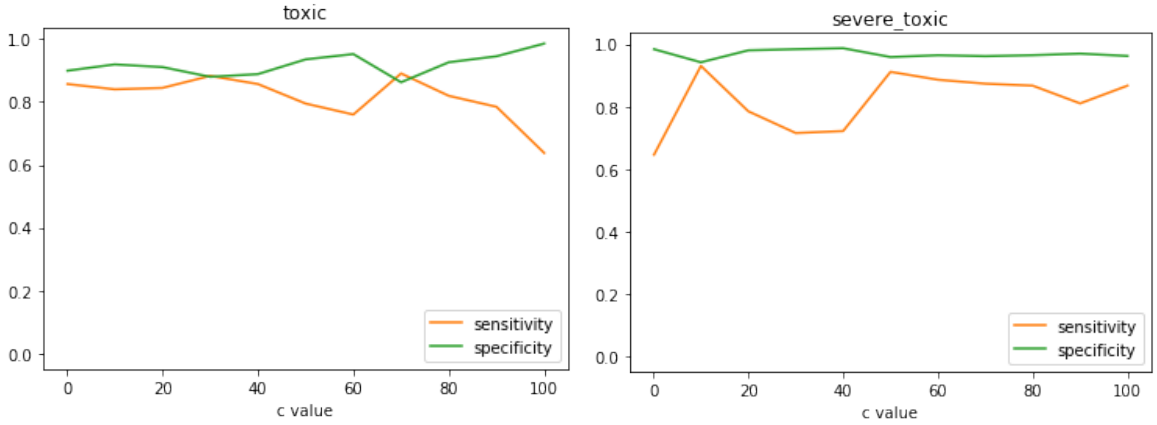
Other parameters:

multi_class set to 'multinomial' instead of the default 'ovr': doesn't seem to improve the performance, not much difference.

e.g. solver='sag', class_weight='balanced', multi_class='multinomial'



e.g. solver='saga', class_weight='balanced', multi_class='multinomial'



penalty set to 'l1' instead of the default 'l2 (when using saga)': performance seems worse and it's very slow. (Didn't draw graphs because it takes too long to draw even one graph).

warm_start set to True instead of the default False: produce similar or worse performance.

Other parameters are not relevant or not applicable. Let me know if I missed out anything!

So the most important parameter is class_weight, other parameters (including the type of solver) don't seem to affect the performance too much.

As you can see there's not really a trend in the graphs, so I'm not sure how we should choose the c-value.

Two other graphs just for comparison:
Sensitivity is much lower when class_weight is set to default. Can't reach 0.8.

Solver='sag', default value for other parameters

Toxic: