Task 1 Results

If we plane to use our model for public alerts, we should use f1 score metric and precision, because we want to notify people only if we quite sure about our prediction. To describe our base model performance, we can note that it has quite bad f1 score and low precision. For better results, we should tune our model base on weighted f1 metric, where we should give more penalty to FP predictions. Base on first 3 FP tweets, we can see that they are absolutely not like tweets about disasters. They have some keywords like ablaze and accidents, but note in context of disaster. For real use of these model, we should have some approve layer. So we can use model results as trigger and then check them by human, that will approve public notification. With 0.77 f1 score and no such big number of disasters in one time, this can work quite well.

FP tweets:

- On plus side LOOK AT THE SKY LAST NIGHT IT WAS ABLAZE http://t.co/qqsmshaJ3N #stlouis #caraccidentlawyer Speeding Among
- Top Causes of Teen Accidents https://t.co/k4zoMOF319 https://t.co/S2kXVM0cBA Car Accident tee Û_
- I still have not heard Church Leaders of Kenya coming forward to comment on the accident issue and disciplinary measures#ArrestPastorNganga