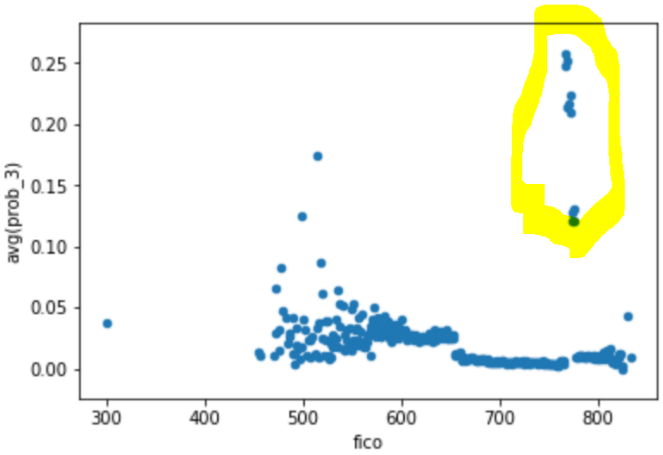
Name: Rohit Garla Net ID: rg3365

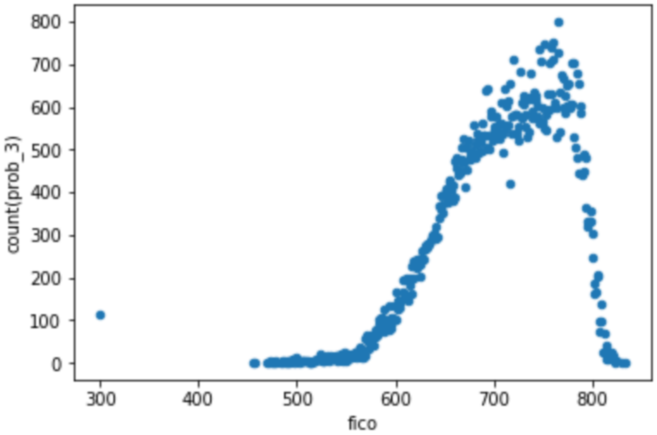
Name: Mengrui Zhang Net ID: mz2258

Week 6 Report

In this homework, we find create some wired sample set to damage the prediction of our model by maximizing the probability 3. In this case, we try to add some poison sample, which contains only label 3 points, to our training set. In order to increase the probability 3 and avoid to being caught.

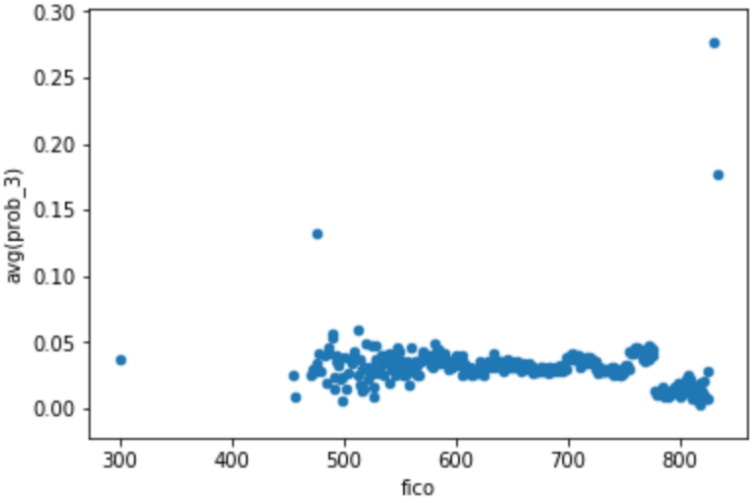
Basically we think, you can not tremendously increase average probability 3 at some given point, which is very obvious, like the following figures.

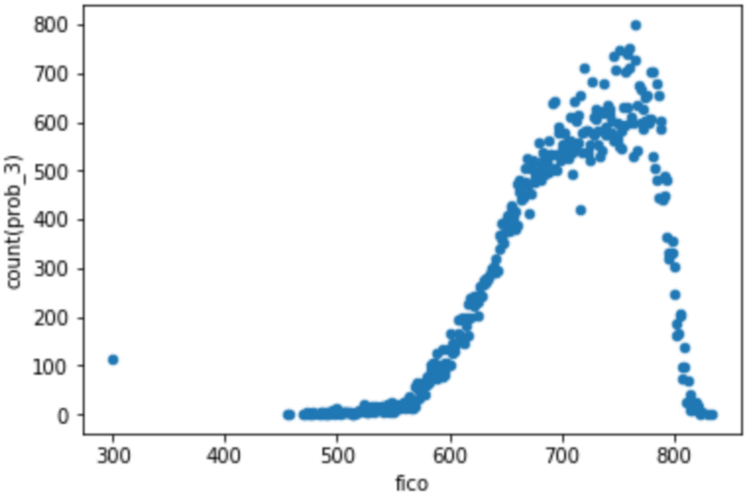




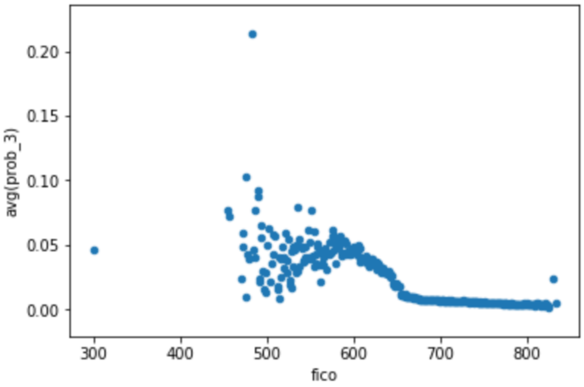
From the above graph, the yellow circle shows that your average probability 3 sharply increase at this interval, so that you may be caught. As a result, we want to make a somehow smooth average probability 3 line.

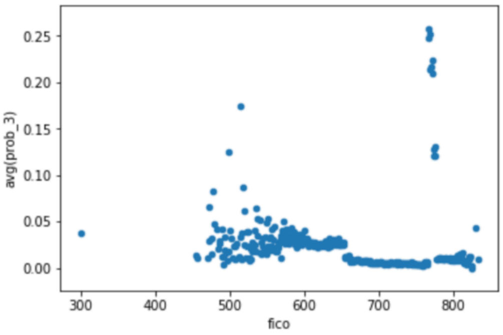
By add some modification, we try to finish the following graph. Though the pattern is a little be different, you get a smooth curve. As a result, it is not very suspicious.





Comparing the original result, and our result, you may say the decrease trend disappears, but comparing to our intermediate result, it's more reasonable. The graph is showing below with left: original, right: intermediate, bottom: new.





Th reason why an opponent would rely on graphs, since graph very directly shows your data set. Opponent will detect whether there are some sharp changes for continued points which is not make sense under our market. In another word, graph is a very good tool to check continuity.

And we choose random forest as our model. If we choose logistic regression as our model. I think our poison set will not make a huge influence. Logistic regression somehow tends to underfit result and it not very sensitive to those outliers, so that you may harder to do this job for logistic regression. In other word, the KL divergence is some small for logistic regression.

And we create two new random forest models with more depth or more trees. You may find that under same poisoned set, new models return lower probability 3, so that we conclude that it's harder to do the same job since you increase accuracy by increasing depth and trees, and random forest is very power to avoid overfitting. But in general, random forest model should be easier to damage than logistic model. Finally small KL divergence model is not enough to avoid detection, since KL divergence is general condition on distribution, however, you may have some extreme points.