5.10

Intro

In this exercise we will discover the general decision rule for the binary classification with arbitrary FP and FN losses. Before we start, I want to point out that this exercise has a wrong formulation. I will first show how to arrive at the correct result. Then, I will argue why the author decision rule don't make sense. Finally, I want to give a shoutout to Dongcheng He, because he was the one that asked me to solve this question, based on his doubts in section 5.7 (Bayesian decision theory). I imagine that like him, there is a lot of people that were confused in this section due to this error in the formula.

Discussion

If you want to express your opinion about this exercise, please create an new issue on my github page, so everyone can contribute.

Solution

We will start from the final result of section 5.7 (The false positive vs false negative tradeoff). There, the author shows that we should pick $\hat{y} = 1 \iff \frac{p(y=1|x)}{p(y=0|x)} > \frac{L_{FP}}{L_{FN}}$. Since this is a binary classification problem we can rewrite $p(y = 1|x)$ as $1-p(y = 0|x)$. Using this property and the fact that $L_{FN} = cL_{FP}$, we have:

$$\begin{aligned}
&\frac{1 - p(y = 0|x)}{p(y = 0|x)} > \frac{1}{c} \\
&c(1 - p(y = 0|x)) > p(y = 0|x) \\
&\frac{c}{c+1} > p(y = 0|x)
\end{aligned} \tag{1}$$

Therefore, we should choose $\hat{y} = 1 \iff p(y = 0|x) < \frac{c}{c+1}$.

Conclusion

In this exercise, we discovered the general decision rule for the binary classification with arbitrary FP and FN losses. The result was not exaclty the same as the author suggest both in section 5.7 (The false positive vs false negative tradeoff) and in this exercise. Now, we will discuss why his rule is incorrect.

Assume the author was right and we should pick $\hat{y} = 1 \iff \frac{p(y=1|x)}{p(y=0|x)} > \tau = \frac{c}{c+1}$. Now assume $c \to 0$. Therefore, the decision rule becomes $\hat{y} = 1 \iff p(y = 1|x) > 0$. This is quite strange, because we have $L_{FN} = cL_{FP} = 0$. In another words, the rule always picks 1, despite the choice of picking 0 having absolutly no cost. An even more bizarre behaviour occurs if we assume $c \to \infty$. In this case, the rule becomes $\hat{y} = 1 \iff \frac{p(y=1|x)}{p(y=0|x)} > 1 \iff p(y = 1|x) > \frac{1}{2}$. In another words, we should pick 1 if the probability of the positve case is greater than the negative case. This says that we could choose the negative case, despite its cost being infinite.

If you perform the same analysis on the rule shown in the Solution section, you will see that the result are consistent. If you think about it, our rule makes much more sense. c encodes how much worse L_{FN} is than L_{FP}. Therefore, it makes sense that as we increase c, we favour picking 1. For big values of c, the rule $\hat{y} = 1 \iff p(y = 0|x) < \frac{c}{c+1}$ says something like: unless you are 100% certain of 0, you should always pick 1. Small values of c have an opposite effect, but within the same line of reasoning.

Overall, the message the rule encodes is quite simple: choices with great costs are very risky. Thus, we should only make them if we have a lot of certanty about it.