Report on Multiclass CWoLa

Duc Nguyen

November 17, 2024

Note that we drew sample A from Gaussian distribution $(\mu, \sigma) = (0, 1)$, B from $(-1, 1)$ and C from $(1, 1)$. Our network is simple, containing 2 dense layers with 32 nodes, ReLU activation, and a dense layer containing 3 nodes for output with softmax activation.

1 Dominated case

In this case, each mixed samples contain a dominant signal, in the ratio of 60-20-20. We achieved high accuracy on both One-versus-All and One-versus-Rest strategies (2), with scores around 82%.

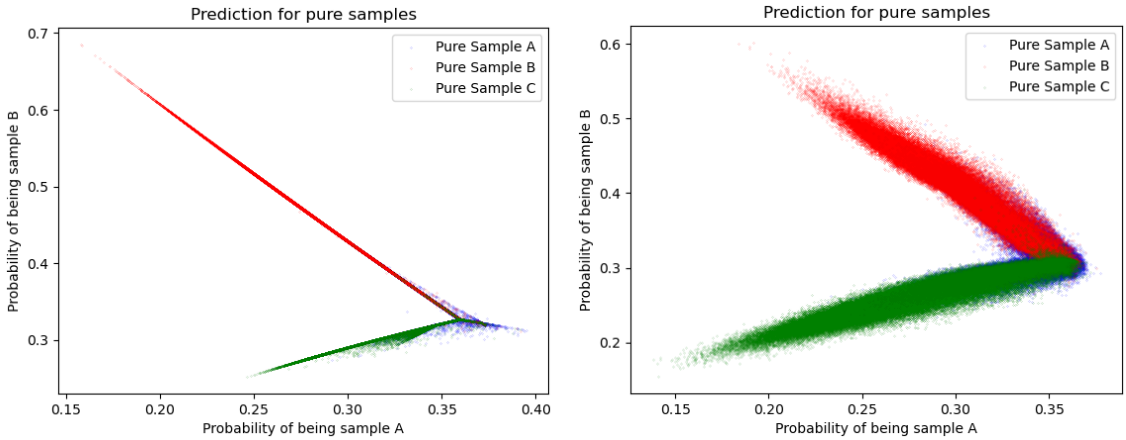

Figure 1: One-hot vector output for pure sample A, B and C in x and y dimension

2 Dominated background, two clear signal

In this case, the ratio is 80-10-10, 70-20-10, and 70-10-20, which means the first one is background A dominant, second one signal B dominant, and third one signal C dominant. The result is similar to the dominated case.

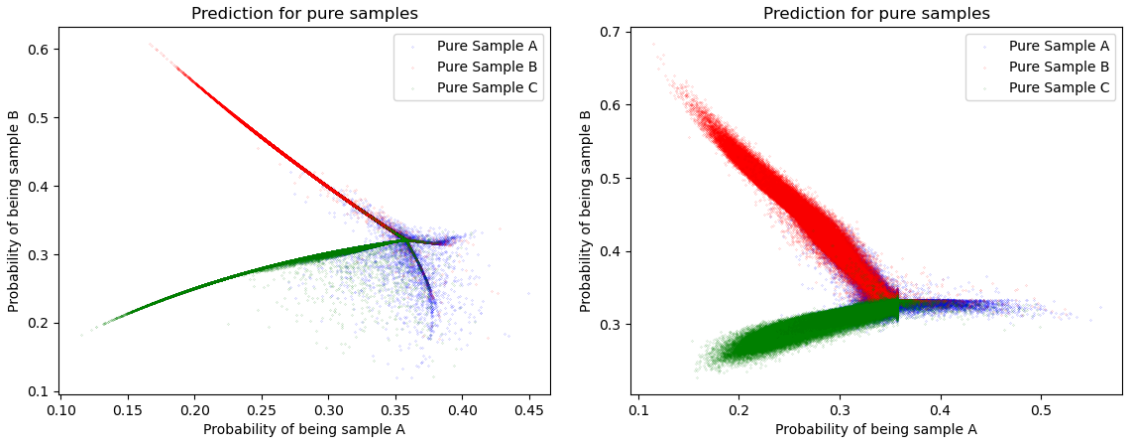

Figure 2: One-hot vector output for pure sample A, B and C in x and y dimension - result for the pure signals case

3 Ambiguous cases

First we begin by examining the case where ratio is $(8, 1, 1)$, $(7, 2, 1)$ and $(6, 2.5, 1.5)$. We found flat accuracy curve (3) for this case when training for this case for the OvR approach. In this case, when tested with pure samples, sample B yielded lowest accuracy. This is

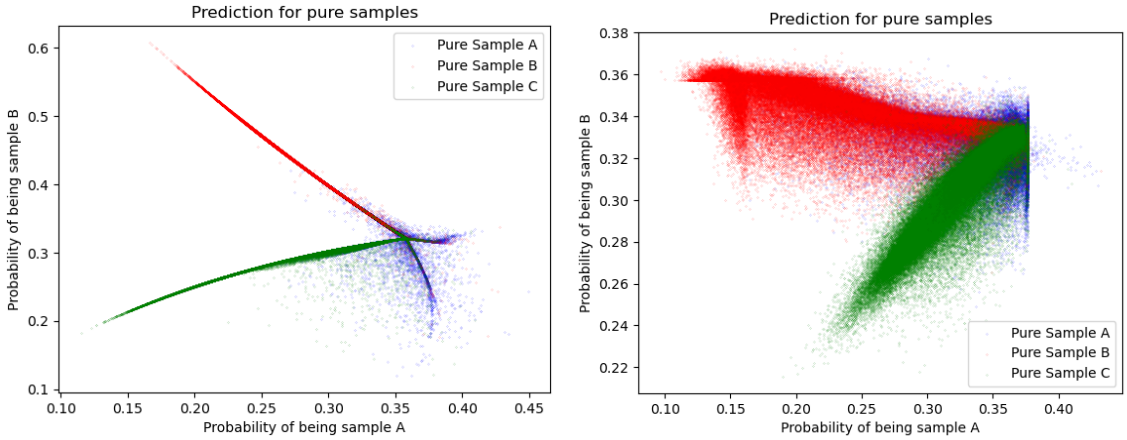

Figure 3: Result for the dominated background, two clear signal case

because most of sample B is being classified as sample C by the neural network in the OvA approach. There's no hope in looking at the mixed data and determine for ourselves which data is which either, as shown in the mixed predictions of 4

Next we examine the case where ratio is $(8, 1, 1)$, $(7.5, 1.5, 1)$ and $(5.5, 2, 2.5)$. Again, similar issues arise, where sample B is mostly misclassified as sample C (5).

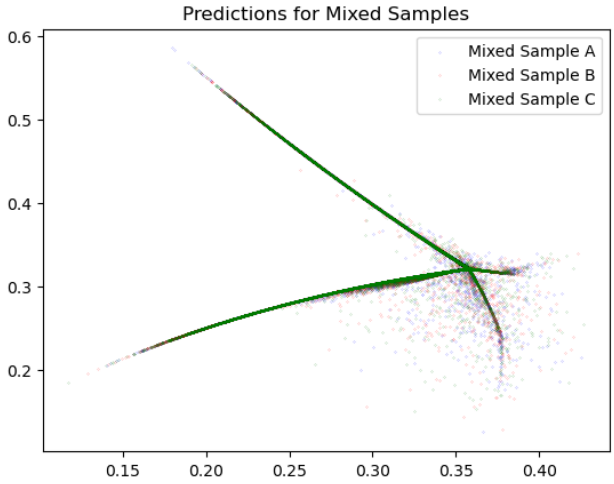

Figure 4: Mixed sample result for the first ambiguous case

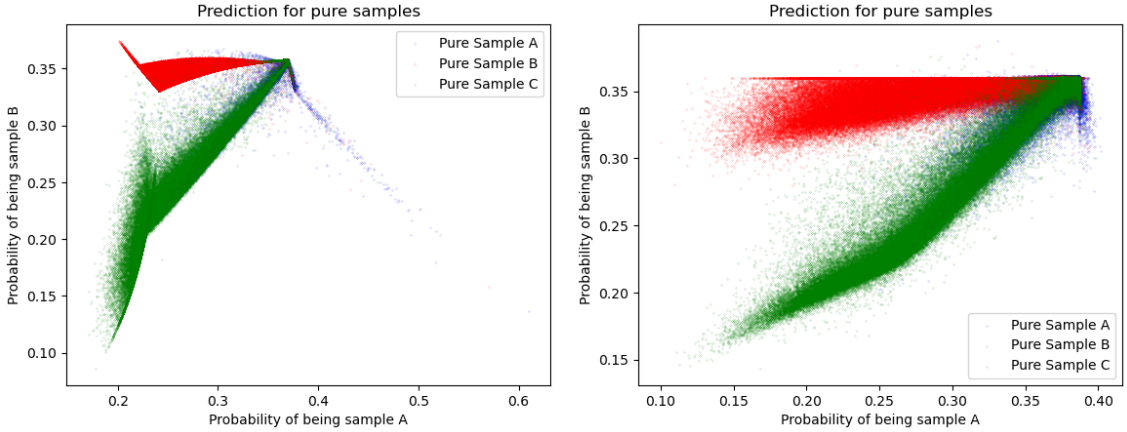

Figure 5: Results for second ambiguous case

4 Bootstrapping

Finally, we can try bootstrapping. Here we will work on the second scenario, training 5 times, 10 epochs each. Bootstrap for this time does improve a little bit, raising the accuracy to $57 - 58\%$ compared to $54 - 55\%$ and removing the fluctuations in accuracy.

5 Final thoughts

Since CWoLa trains on mixed distribution, the ultimate factor that this method relies on to distinguish the pure samples is the difference between the mixed samples and whether or not it maps nicely to the difference between pure samples. Therefore, we need to find ways

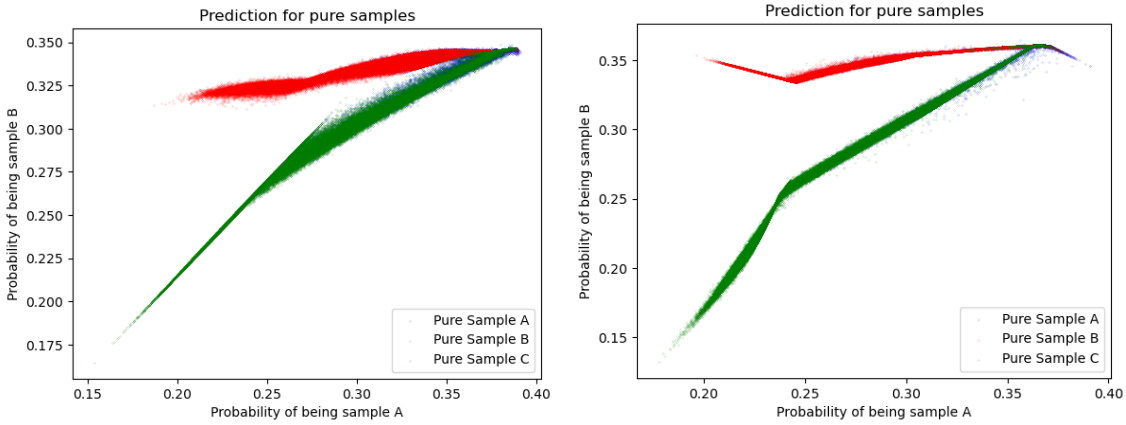

Figure 6: Results for bootstrapping (left) and non-bootstrapping (right) for ambiguous scenario