

Machine Learning - Assignment 2

Submitted by: Md. Tanvir Jawad (tj2303@bard.edu)

ANALYSIS & DISCUSSIONS:

- What makes the model “good”?

A good program is one that doesn't make mistakes, especially not saying a dangerous mushroom is safe. We look at a few things to decide if our program is doing a great job:

Precision: This means when our program says a mushroom is safe, it's really safe. We really care about this because we don't want to eat something dangerous by mistake.

Recall: This means our program is good at finding all the safe mushrooms and not missing any.

Accuracy: doesn't tell the whole story, especially if most mushrooms are safe anyway.

We look at how often it's right, especially making sure it doesn't say a dangerous mushroom is safe. This is important because we don't want to get sick from eating the wrong thing.

- **Wow similar are the two classifiers in terms of how they perform on the test set and how they weigh different features. If they are very similar, why is this the case? If not, why not?**

The perceptron and logistic regression are two methods used for classifying items, such as determining whether a mushroom is safe to eat. They each have their own way of dealing with information and making decisions.

The perceptron method tries to separate different categories by drawing a straight line. If the data can't be neatly divided by this line, the perceptron might not perform well. In contrast, logistic regression estimates the probability that something belongs to a certain category. This approach can be more flexible, especially when the data is complex or doesn't line up neatly.

Both methods can also tell us which characteristics are important for making classifications. However, they use this information in different ways. The perceptron combines these characteristics in a simple mathematical way to make a decision. Logistic regression, on the other hand, uses them to calculate how likely it is that something falls into a particular category. When both methods agree on the importance of certain characteristics, those are likely very reliable indicators. If they disagree, it could be because of their distinct ways of interpreting and handling the data.

- **Finally, which of your two classifiers is better and why?**

If both of our computer programs show similar results, we need to dig a bit deeper to pick the best one. It's really important that we don't mistakenly call a dangerous mushroom safe, so we pay extra attention to how often the program is right and how many safe mushrooms it can find.

The Perceptron program is pretty straightforward and quick, which is great when we're short on time or our computer isn't that powerful. But, Logistic Regression, another program, takes a bit longer. It doesn't just say yes or no; it also tells us how sure it is about each mushroom. This extra detail can really help us make better choices.

Even though Perceptron is faster and simpler, Logistic Regression often ends up being the better option for figuring out which mushrooms are okay to eat. It's better at handling new kinds of mushrooms it hasn't seen before, gives us a better understanding of why it thinks a mushroom is safe or not, and tells us how confident it is in its guess. All these things make Logistic Regression not just more powerful but also more trustworthy for deciding which mushrooms we can safely eat.