Olivia Samples

Case Studies

Lipscomb University

September 2020

Data Science for Business: Visualizing Model Performance

Model performance can be based on multiple different metrics, but choosing the right one for the right situation can be more beneficial than the best performing metric. For instance, the most well-known metric for model interpretation is accuracy. Accuracy tells us how many instances the model predicted correctly versus the total number of instances. For example, accuracy tells us that 95% of the votes are counted accurately, but that leaves room for 5% of the votes to go to one side or the other. For tight races, it is important to know which side the inaccurate solutions might come on, or if there will be more false positives or false negatives. When these ratios are important to you, AUC is a better metric.

Another metric that can be useful is a Profit Curve. Profit curves depend on the proportion of positive and negative instances in the population along with the cost and benefits of the instances. When these two factors are available, profit curves can serve as a good metric. More likely, when these factors are unknown, it is better to use an ROC Curve.

In both of these situations, AUC and ROC Curve tend to be better solutions as they are dependent on each other. A ROC (Receiver Operator Characteristics) Curve accommodates uncertainty by factoring all possibilities, and AUC (Area Under Curve) represents this accommodation within a single statistic. While these metrics are extremely useful, they are difficult to digest. A valuable alternative is a Lift Curve.

A lot of times, data science is all about business. It doesn’t matter if your model performs very well if you can’t explain how it works to the business admins responsible for its implementation. Where accuracy and profit curves are intuitive, ROC and AUC are not. On the other hand, the Lift Curve offers a stable middle ground. It offers an understandable concept of how much better your model is performing than randomness would have. It does so by sorting the true positives by an effective ranking classifier, then essentially dividing the top half by the bottom half resulting in a result greater than 1 if the model performs well.

With this metric, you are able to relay to your business counterparts how much better your model performed than before. This gives them an exact metric to understand why or why not they need to implement the model, which is all they really want. All in all, it is better to use the most simple and truthful metric for expressing model performance, where lift curves can be some of the best options.