Looking at Measurement and How We Evaluate the Impact of Reinforcement Learning Over Traditional Predictive Analytics

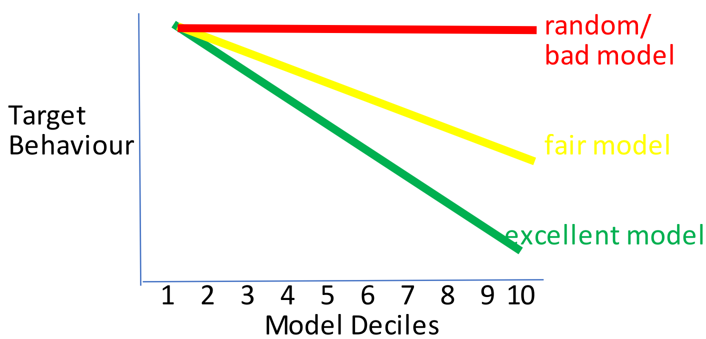
Written by Richard Boire at Boire Analytics

Are we really at a new stage of so-called industrial development much like how the car replaced the horse and buggy as the main mode of consumer transportation? This would appear to be the case by some thought leaders and practitioner leaders that reinforcement learning (RL) will and should replace traditional prediction analytics as the mode of targeting.

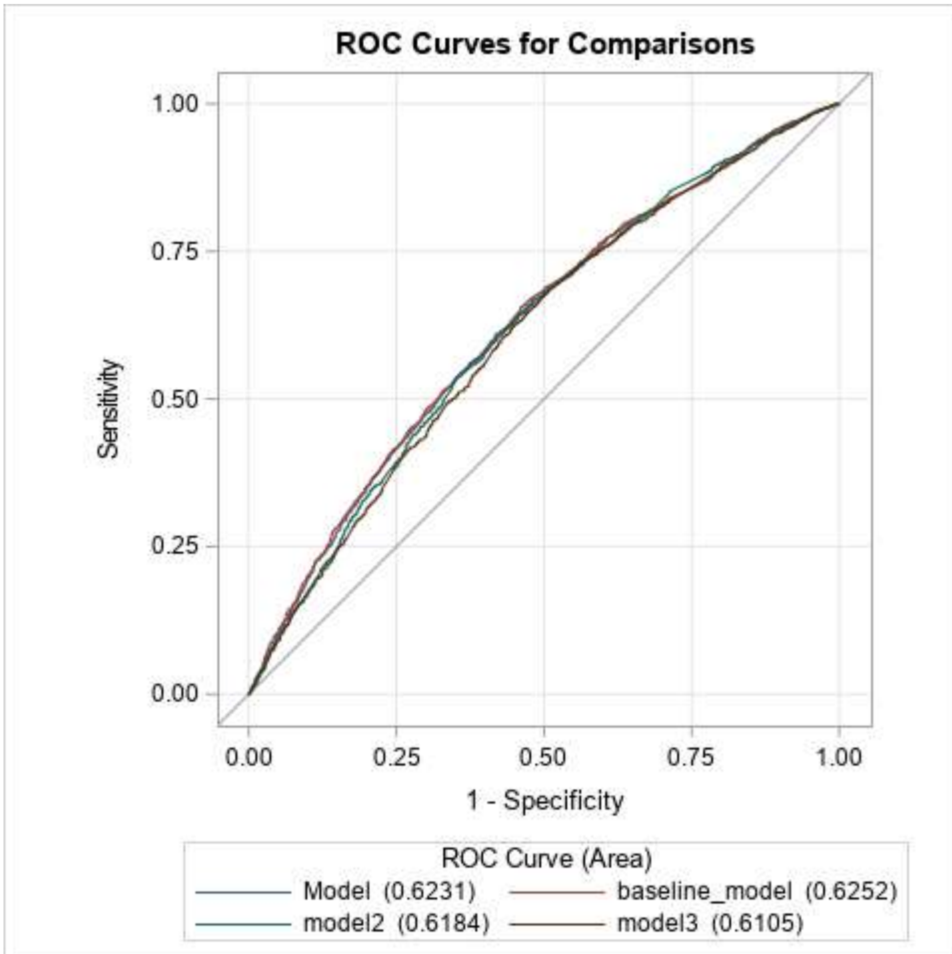
With my many years as a data science practitioner (30 years+), I have been trained in the discipline of measurement in evaluating what works vs. what does not work. In assessing the value of predictive models, the goal was to always be able to measure incremental benefit or improved performance. So, the same philosophy and principles should apply to RL. But let's first spend some time talking about measurement.

Looking at Measurement

With the advent of software that provides an AML (automated machine learning) environment, the measurement process in evaluating predictive models has been greatly facilitated. As discussed in many previous articles, the use of decile charts/gains charts can be produced to determine the predictive model that yields the best rank ordering of the observed target behaviour or steepness of the slope when graphing the target behaviour across the model deciles.



Another perspective, more common in the academic data science textbooks in data science, is looking at the AUC or ROC curve where the area between the parabola and the straight line represents the power of the model. See below.



In this example, we are measuring the % of the target behaviour on the y-axis while % of records are ranked ordered by model score and are plotted on the x-axis. The straight line indicates random as the top 10% of scored records yield 10% of the observed target behaviour, the top 25% of scored records yields 25% of the observed target behaviour, and so on. Meanwhile, the baseline model is the best model (.6252 score) which indicates that the top 25% of scored records yields approximately 40% of the observed target behaviour since no other model has a higher score.

Every model and its desired outcome can be evaluated from either of these two perspectives which is offered by all the industry modelling software leaders. As stated above, practitioners can evaluate a variety of different options or even combine them through the use of ensemble modelling which has been discussed at length by many of the leading-edge practitioners. The advantages of this type of software is that we can very quickly develop and compare the performance of many models instantaneously. So, how do we apply this to reinforcement learning (RL)?

Defining Reinforcement Learning and some of its applications

But first what is reinforcement learning. The concept of reinforcement learning is that you are conducting activities against steady or existing states in order to move to a desired state and receive a reward. Reinforcement learning is not new as the mathematics of Markov Models were its earliest applications of assessing steady states and the activities required to achieve a desired state and a reward. With the advent of AI and deep learning, reinforcement learning is now using the concepts of RNN(recurrent neural nets) and specifically LSTM(Long-Short-Term Memory) where the network places more weights on its more recent results but with the intention of optimizing that reward outcome.

The benefits of RL can be enormous and are being used in applications such as robotics and more specifically self-driving cars. But in the area of consumer behaviour which has been my area of career focus, can we observe benefits which are incremental to traditional predictive analytics and what might be those specific business applications. Let's explore this a little more closely. If the ultimate objective or reward in consumer behaviour is maximization of customer profitability, then what are the activities that achieve this goal. In other words, can we use RL to maximize customer profitability? As stated before, most of the work in building any advanced analytics algorithm is going to be in creating the right information or analytical platform. The organization of this platform can be quite complex especially when there are multiple metrics comprising customer profitability. Think of a bank which generates profitability from a variety of metrics (deposit, risk, lending, investment, etc.). And think of the many desired states that may optimize each metric. At the end of the day, though, the end deliverable is predicting the right activity in order to achieve the desired state of customer profitability maximization. Yet, could a more traditional approach be considered? The rationale behind this thinking is that in the realm of predicting consumer behaviour, there is much randomness/noise (unexplained variation). For example, the buying behaviour of consumers is going to exhibit more randomness or noise than how a car should navigate on a given road. It is this level of randomness where often the more simpler type techniques will suffice in delivering the same level or even better level of model performance than more advanced deep learning RL algorithms. Why? Deep learning techniques need a strong signal to noise ratio which is the case for image recognition and NLP and self-driving algorithms. Yet, for consumer behaviour, it is not necessarily a slam dump. But this does not address the question of how we might evaluate RL vs. traditional predictive analytics techniques. Again, we will set up a measurement framework. However, the framework provides a longer-term perspective as we need to really compare multiple models and their predicted activities versus the RL outcomes over a period of time.

Creating the Measurement Framework for RL and Traditional Predictive Analytics

One approach would be to set up a random sample of a customer database versus the balance of customers. In both the random sample and the balance, essentially, we are trying to determine the next best action which will ultimately optimize customer profitability. In the marketing literature, much has been written about next best action as the ultimate marketing goal. Of course, in setting up the measurement environment, these series of actions would have to be defined in both datasets. But once these actions are determined, it is the selection of actions for a given customer that will differ between RL and traditional predictive analytics. In RL, the action is determined based on the output of the deep learning network as it tries to optimize the reward outcome. In traditional predictive analytics, separate predictive models would be built for each action. Using some sort of score normalization that looks at expected value, the selected action for that customer would be dictated by the model yielding the highest expected value. Again, the heavy work would be in creating the analytical environment which is similar for both RL and traditional predictive analytics. Once the analytical environment is established, tools are available to quickly create and update the many predictive models that would be required. The use of Auto ML tools to constantly develop new models as well as update them would simulate the automated black box environment of deep learning. Yet, one current advantage at least so far in traditional predictive analytics is the ability to explain how the models are working. But besides model explainability and how quickly we can select the desired action, the most important factor in comparing the two approaches is performance. In this type of measurement environment, we would monitor both approaches as follows:

Approach	Average Customer profitability 1 month	Average Customer profitability 2 month	Average Customer profitability 3 month	….	Average Customer profitability 12 months
Traditional Predictive Analytics					
Rinforcement Learning (RL)					

This would be the report card in really assessing whether RL can work in a consumer behaviour type market. Such an approach is not really new as one bank that I worked with in the mid 90's was trying to determine the benefits of using predictive analytics in their decision-making as opposed to the status quo which consisted of business rules based on post analysis. They adopted the same approach of setting aside a random sample for predictive analytics activities and the balance for the status quo. Given the growth of predictive analytics in the last 25 years, it is not surprising that predictive analytics emerged the winner.

Emerging technologies are all around us and it is easy to become consumed by the Kool-Aid of not looking like a Luddite. But sound measurement should always be the foundation of new initiatives and activities that we decide to undertake. Many of the experienced data science and analytics practitioners were schooled in this by those company pioneers of data science and analytics. I think it is time to adopt this learned philosophy towards all emerging technologies including RL.