Article from

Predictive Analytics and Futurism

April 2018
Issue 17

Better Tools—
Less Dukkha

By Dave Snell

Many of us from the Western world are not familiar with the word "dukkha." Like many words of foreign origin, it does not have a one-word English translation. According to Wikipedia, it is the first of the Four Noble Truths of Buddhism. It is also found in scriptures of Hinduism, and it refers to "the fundamental unsatisfactoriness and painfulness of mundane life."[1]

What does this have to do with predictive analytics and futurism, and the associated techniques we embrace in this section? Perhaps this new era of artificial intelligence (AI) and machine learning will help liberate many of us from the dukkha of our current routines. Many of us spend boring hours commuting to and from work—sometimes in a paradoxical situation where as drivers we must patiently wait in long lines of traffic, yet we must be constantly vigilant to avoid accidents—often induced by the boredom of the waits. Autonomous cars may not only increase our safety, but also permit us to luxuriate in creative thought, having delegated the tedium of traffic mindfulness to our vehicles. We spend far too much time at work (and at home) on repetitive tasks that become mind-numbing rather than mind-expanding.

Some might argue that autonomous AI should never be trusted for life-critical decisions. I, for one, am ready and willing to delegate many of the processes and decisions of the day to AI, just as I delegate the life-critical tasks of breathing and digestion to my autonomous nervous system. In fact, I cannot imagine how tedious and stressful it would be to have to remember to breathe in and breathe out thousands of times per day; or to consciously have to tell my heart when to contract the left ventricle and send essential oxygenated blood to each of my cells.

Throughout human history, we have developed tools to do the "heavy lifting" for us: from shovels through backhoes. On the data assimilation, number-crunching and presentation side, we also are improving our tools, and this issue has several articles about new tools and techniques that can help you reduce your dukkha:

- Starting with Anders Larson in his "Chairperson's Corner," we are reminded of the importance of upgrading your tool set. Sure, there are sometimes temptations to treat every problem as a nail for your new hammer; but as Anders says, "just because everything isn't a nail, that doesn't mean there aren't nails out there that you've been hitting with a spoon." He talks about tools such as random forests that improved his analysis of the impact of multidimensional factors on health costs; and he even describes an upgrade to some of his Excel workbooks by using a function that computes vector products on a conditional basis.

- Next, "Parallel Cloud Computing: Making Massive Actuarial Risk Analysis Possible," by Joe Long and Dan McCurley, walks us through a cloud use case where they were able to cut a three-month machine learning exploration project down to just under four days using a mixture of open source tools and a cloud environment. That freed up a lot of time for them to digest the results, and run variations that would not have been feasible with a single processor approach to the project. Yes, they had to spend some time on the learning curve for parallelization; but it resulted in much faster throughput. Supposedly, Abraham Lincoln said, "Give me six hours to chop down a tree and I will spend the first four sharpening the axe." Lincoln would surely have viewed a 25-fold efficiency return as a wise investment.

- Moving on in our description of new tools, Michael Niemerg tells us about a novel technique. "The Forgery Game:

Generative Adversarial Networks" describes a generative adversarial network, or GAN. This is a very recent technique in artificial intelligence algorithms—introduced in 2014. A GAN is an unsupervised machine learning technique and can accomplish some interesting, and perhaps disturbing, outputs. Basically, models compete with each other and generate synthetic data. In one type of application, the result can be indistinguishable from a real photographic image (thus, the forgery game). This is leading-edge stuff; and as I am writing this issue introduction (Chinese New Year—新年快乐), I see an article about GANs in use to analyze molecular genetic mechanisms to create new synthetic drugs.[2]

- Not every tool has to be new, of course. Some are older ones that just have been underutilized. Ben Wolzenski led our "Blue Ocean" Delphi study back in 2009; and it predicted some nearly heretical ideas back then, such as pet insurance and custom-designed coverage developed online. Now, they have become important products with rapid growth. In "Why Consider a Delphi Study?" Ben describes advantages of this largely qualitative rather than quantitative approach to forecasting. It can provide value when other methods cannot, and can also serve as a second opinion for the other methods. He also details how our section has provided leadership in this technique in previous studies, and mentions another SOA Delphi study being launched now.

- I wrote the article "Hierarchical Clustering—A Recommendation From a Nonhierarchical Manager," where I describe a bottom-up, or agglomerative, technique that is more visually appealing to nonmathematicians than the more common k-means approach to clustering. Sometimes we overlook the fact that most senior managers are not actuaries or data scientists; and a tree-like visual that shows both the natural groupings you have discerned and the relative dissimilarity among the various groups, for even a multidimensional set of groupings, might be easier to understand, and thus more likely to be accepted.

- As the amount and types of data continue to increase, the complexity of models can be a limiting factor in their utility. Jeff Heaton, in his article, "Feature Importance in Supervised Training," addresses the issue of choosing which factors are the more important ones. Jeff takes us through model-specific feature ranking, model-agnostic feature ranking, and multivariate feature ranking. Removing unimportant features can increase both the speed and the accuracy of your models. This is especially important when you are employing feature engineering, which can benefit from feature importance evaluation to reduce the number of combinations involved in pair-wise multivariate considerations.

- It is nice that we have these new modeling tools available to us; but how do we share them with the folks who do not have, or even want to have, RStudio or a predictive analytics toolbox on their PCs? What if they want to get insights from your modeling efforts; but they do not wish to have to write R code to do that? In "Shiny: Another Step Forward in Data Democratization," Eileen Burns introduces us to a tool that addresses that concern. It's called Shiny, and the name is apt because it allows you to create an attractive and intuitive web application where non-programmers can try out your model and be creative and productive with it. Shiny can help you share your R apps with a larger base. The example she describes for us is a project where she put a web front end on the new *PAF Newsletter* Catalogue.

Eileen's article is also my segue to a new feature you should all be enjoying now—the newsletter index of all 195 articles from our section newsletters. These go back to September 2009, when the old Futurism Section became the Forecasting & Futurism Section (later Predictive Analytics and Futurism . . . as a result of a Delphi study). We know that most actuaries love Excel, and especially like to filter and sort and do lots of other data manipulations with it. On the newsletter webpage, you can now download your own copy of an Excel workbook with several columns for each article. If you want an actuarial perspective on agent-based modeling, neuroevolution of augmenting topologies (NEAT), hidden Markov models, genetic algorithms, or dozens of other topics, you can search, sort and filter it as much as you wish. Best of all, when you find the article you want to see, you can click on the hyperlink and go right to that issue! Thanks to Nick Hanewinckel, the PAF Section Council, section specialist Jessica Boyke and staff partner Beth Bernardi, we all have a handy new research tool.

Perhaps we can't completely escape dukkha; but the tools and techniques described in this issue ought to make it less mundane and less painful for you.

Enjoy! ■

Dave Snell, ASA, ACS, ARA, ChFC, CLU, FALU, FLMI, MAAA, MCP, teaches AI Machine Learning at Maryville University in St. Louis. He can be reached at *dave@ActuariesAndTechnology.com*.

ENDNOTES

1 Quoted from *https://en.wikipedia.org/wiki/Dukkha* (accessed Feb. 16, 2018).

2 *http://www.mauldineconomics.com/tech/tech-digest/right-to-try-our-best-shot-at-saving-healthcare*