Article from

Predictive Analytics and Futurism

December 2016
Issue 14

From the Editor: Insights from a Dead Salmon!

By Dave Snell and Kevin Jones

Our SOA Cultivate Opportunities Team, and the PAF section have been saying for a long time that actuaries provide some valuable insurance and risk management knowledge that some of the other quants (MBAs, CFAs, CPAs, … substitute you're favorite xxAs), competing on the predictive analytics fronts with us cannot contribute. We feel that this is an important differentiator.

At last, the media has described the folly of analytics without subject matter expertise: Apparently, tens of thousands of studies have been published that cite fMRIs (functional MRI scans) as their justification; yet the underlying software for analysis of fMRIs had a software glitch that may have created false positives up to 70 percent of the time.

"A graduate student conducted an fM.R.I. scan of a dead salmon and found neural activity in its brain when it was shown photographs of humans in social situations. Again, it was a salmon. And it was dead."

http://www.nytimes.com/2016/08/28/opinion/sunday/do-you-believe-in-god-or-is-that-a-software-glitch.html?action=click&pgtype=Homepage&clickSource=story-heading&module=opinion-c-col-right-region&region=opinion-c-col-right-region&WT.nav=opinion-c-col-right-region&_r=0

This issue offers you a collection of basic and leading edge predictive analytics. It also gives you insurance applications and some insights into how to employ the analytics in a more useful manner—one that minimizes the chance of dead salmon results.

"Chairperson's Corner: On Volunteering, Learning, and a Sense of Community," by Ricky Trachtman: Our incoming chairperson, Ricky, writes about a common theme throughout the PAF section: community. This may be viewed as a temporal update to the "it takes a village" concept. Ricky tells how the PAF community (then F&F—forecasting & futurism) helped him get involved, and how the volunteer spirit of our community becomes contagious—in a very good way. Yes, you put in a lot of time without direct compensation; but the result is highly rewarding!

"Looking Back and Ahead," by Brian Holland: Brian gives us a reminder of the many cool accomplishments of our section just in the last year. It is an impressive list of results from our tireless and enthusiastic volunteers. It is extra work to volunteer your time and expertise; but the rewards are well worth it for others and for yourself. As Brian says, "A lot of the SOA is what we make of it, and the sections are a great way to pitch in and make the SOA what you want it to be."

"Making Predictive Analytics Our Own," by Joan C. Barrett: The Society of Actuaries is embracing predictive analytics enthusiastically now; and our board partner, Joan, provides sage advice on how we can take back much of the predictive analytics space in the financial services industries. She tells us how she personally shows that actuaries add value beyond the mathematical mechanics of modeling. She stresses monitoring experience (with several cogent examples) and employing behavioral economics (again, with examples) to check the sensibility of model results, and improve the models.

"Deciding What to Research: How to Spot and Avoid Bias," by Kurt Wrobel: Kurt provides an article without even one mathematical formula in it; yet it contains essential advice for any actuary involved in predictive modeling. He explains how studies can become biased; and how you can watch for and avoid or minimize

CONTINUED ON PAGE 5

bias in your work. His tips can serve as a guideline for making sure you have objectivity, quality, and relevance in your modeling work.

"Five Myths and Facts about Artificial Intelligence," by Anand Rao: Movies and television have perpetuated many misconceptions about Artificial Intelligence in the name of drama, but Anand is stepping forward to clear up a few of these. In order to see its future in the insurance industry, one has to understand what AI is, what it can do, and what its relationship will be to human beings. It may not produce magical solutions to all problems, but has great potential to enhance decisions.

"Abstractions and Working Effectively Alongside Artificial Intellects," by Dodzi Attimu and Bryon Robidoux: In any technical field, abstraction is necessary to communicate ideas without getting bogged down in details. Dodzi and Bryon outline different types of abstraction and explore their uses in a software setting. They push this even further into layers of abstraction in machine learning. Remarkably, the human brain's approach of Sparse Distributed Representations (SDRs) can be used to create a robust learning solution.

"Machine Learning: An Analytical Invitation to Actuaries," by Syed Danish Ali: In ratemaking, do we trade specific risk for systematic risk? Machine learning can give new understanding in that question. Danish shows how it can be applied in exploring data, predictive modeling, and unstructured data mining. Moreover, it's important in understanding big data and its effect on the insurance industry.

"Use Tree-based Algorithm for Predictive Modeling in Insurance," by Dihui Lai and Bingfeng Lu: They provide a clear explanation of tree-based models, starting from a simple binary tree, through CART (classification and regression tree), to an ensemble of trees in a random forest, or boosted tree. In addition, they give advice on when to use which type, what their particular strengths are, and when a tree is not the best choice for your model.

"Creating a Useful Training Data Set for Predictive Modeling," by Anders Larson: Data scientists (and actuaries involved in predictive analytics) rant about the value of clean data of sufficient quality and quantity for statistical significance. As Anders says, "This goes beyond the simple 'garbage in, garbage out' principle," but what is actually meant by good data, and how do you obtain it? Read here the key terms that define a good training set; and how to create and use training sets that will provide insights on situations involving transactional data, and expanding populations.

"The Random GLM Algorithm: A Better Ensemble?" by Michael Niemerg: What if you could combine the benefits of a generalized linear model (GLM) with those of a random forest? Michael describes a lesser-known technique that attempts to do this—the random generalized linear model (RGLM). It uses randomization and bagging (often associated with random forests) and applies them to a GLM. He compares the RGLM to random forest results, and provides overall observations about the strengths and weaknesses of this method.

"Collaborative Filtering for Medical Conditions," by Shea Parkes and Ben Copeland: Collaborative filtering systems, known to some of us as recommender systems, are used by Amazon, Netflix, Google, and many others to determine what you might wish to see or buy based on what others like you have seen or bought, and on what other items are similar to ones you already saw or bought. Shea and Ben take us beyond the simple discovery of frequent item sets, and introduce ratings based upon estimated latent factors. Their example for a diabetes patient shows how this approach can improve the coding of patient comorbidities.

"Getting Started with Deep Learning and TensorFlow," by Jeff Heaton: Deep Learning is the machine learning technique of choice for most image recognition and time series predictive analytics models. Jeff names the four powerhouse open source toolkits from Amazon, Baidu, Google, and Microsoft; and then he describes the Google product, TensorFlow, in more depth. TensorFlow, was introduced in 2015. It is new, and hot! Jeff shows us how to use it—with code examples to get started. Follow Jeff's article, based on the graduate course he is teaching at Washington University.

"Guide to Deep Learning," by Syed Danish Ali: Unstructured data presents data scientists with many challenges, but Deep Learning can be a valuable tool for insight here. Here, Danish gives an overview of the purpose and structure of Deep Learning, as well as the challenges. There are also many variants on deep architectures outlined.

"Introduction to Using Graphical Processing Units for Variable Annuity Guarantee Modeling," by Bryon Robidoux: The need for faster processing has run into a bottleneck with traditional central processing units (CPUs) and the graphical processing unit (GPU), where dozens or even thousands of special purpose calculation engines can work in parallel, offers the potential for breakthrough speed improvements. However, it is not as simple as buying and inserting a high-end GPU card. Bryon takes us through a reality check on how, and when, to use these powerful processors effectively. ■

Dave Snell, ASA, MAAA, is technology evangelist at RGA Reinsurance Co. in Chesterfield, Mo. He can be reached at *dave@ActuariesAndTechnology.com.*

Kevin Jones, FSA, CERA, is associate actuary at Milliman in Buffalo Grove, Ill. He can be reached at *Kevin.Jones@Milliman.com.*