



Article from

## **Predictive Analytics and Futurism**

June 2017

Issue 15

# From the Editors: A Rainbow of Opportunities

By Dave Snell and Kevin Jones

**W**elcome! Once again, we have another collection of exciting articles relating to predictive analytics and futurism. Usually, we try to summarize a theme here; but the theme in this issue seems to be “something for everyone,” and we mean that in a positive way. The PAF section has become the fastest-growing one in the SOA, and we are proud to say that we have doubled in size over the last couple of years. We would like to think that the newsletter is one of the factors that attracts and retains members; and your excellent article contributions are showing us the vast variety of interests in both the highly quantitative, sophisticated algorithms of predictive analytics and the less tangible but forward thinking aspects of futurism. Before we describe the articles in this issue, though, we want to point out two new structural features:

1. We moved the issue date from July to June (this makes the two semiannual issues six months apart (June and December), and it ought to help make our pipeline of contributions more consistent between the two issues each year.
2. We will be adding a new feature to the newsletter section of the PAF website <https://www.soa.org/sections/pred-analytics-futurism/pred-analytics-futurism-newsletter/> so that you can download an Excel file of all the previous article titles and descriptions along with hyperlinks to the issues that contain them. This will allow you to sort, filter and do all the creative things actuaries use Excel for to manage your library of more than 150 (and growing) articles of interest.

Now, let’s talk about what we have for you in this issue:

“Chairperson’s Corner: Jump on the PA Bandwagon” by Ricky Trachtman: Ricky claims in his article that English is not his first language, but he is eloquent in describing how to answer his neighbor’s question “Don’t actuaries predict stuff? You should jump on the bandwagon.” Ricky describes how we are both jumping on the PA bandwagon and, in some cases, leading the band.



“Podcasts—A Drill You Can Enjoy!” by Geof Hileman: One area of PAF leadership in predictive analytics has been in the use of podcasts. Geoff describes how you can use these affordable (free), accessible (on your smartphone or similar mobile device) and highly informative audio resources.

“Understanding Autoregressive Model for Time Series as a Deterministic Dynamic System” by Dihui Lai and Bingfeng Lu: In this article, Dihui and Bingfeng describe some of the “art” as well as science in modeling times series data. They introduce a seasonal difference variable—a clever approach that avoids modeling the periodic behavior (more difficult) of these autoregressive models and provides the reader with some guidance for developing an instinct while working with them.

“Predictive Model Building 101” by Dorothy L. Andrews: Dorothy had an ambitious goal for this article—“a guide to help you navigate through 10 modeling phases for building a predictive model and provide you with some insights as to how to overcome obstacles you will likely encounter along the journey.” She did it! This is a thorough description of how to get started, how to do it, and then how to validate, test, integrate and monitor it. It’s worth saving as a checklist for many predictive analytics projects.

“Predictive Modeling Techniques—A Case Study in Resolving Correlated Explanatory Variables” by Vincent J. Granieri: Following up on his article last year that introduced us to using the Cox Proportional Hazards Model in an underwriting environment, Vince shows how regressing data to find the impact on a dependent variable of many explanatory variables is a worthwhile exercise when building an underwriting debit/credit model.

“Ground Assessment of Soft Skills in Actuaries” by Syed Danish Ali: Danish has written several articles for us, and up until now they were all very technical. He certainly has the technical skills, but this issue shows a softer side to his talents. In this article he quotes Nietzsche: “You must have chaos within you to give birth to a dancing star.” We have had several articles on behavioral economics and the importance of the softer, nonquantitative, aspects of predictive analytics, but we think you’ll enjoy his unique perspective on it.

“Using Python to Solve, Simplify, Differentiate and Integrate Mathematical Expressions” by Jeff Heaton: We knew Python was cool. But Jeff introduces us to an amazing free and open-source package that allows you to essentially get many of the benefits of Mathematica (expensive) and even have program control over it. Imagine doing symbolic equation simplifications, differentiation and integration within your code with almost no effort. We recommend it!

“On Building Robust Predictive Models” by Mahmoud Shehadeh: Mahmoud takes a deeper look at statistics for us and clarifies how to use a single hold-out validation on a large, publicly available training set (more than 100,000 records). In his article he shows the dangers of using generalized linear model (GLM) results without further investigation of a larger number of samples.

“Speculative Fiction Contest and the Predictive Analytics and Futurism Section Award” by Ben Wolzenski: In case you ever wondered where futurism fits into our section name, Ben describes the recent actuarial speculative fiction contest and lets you know how to get to more than two dozen original science fiction stories written by actuaries and involving some aspects of predictive analytics. They are entertaining and often quite thought provoking.

“Data Visualization for Model Controls” by Bob Crompton: Predictive analytics best practices are not confined to one SOA section; and Bob shares some from his recent article from the March 2017 issue of the *Financial Reporter*. As Bob says in his article, “Can we do better than subject model reviewers to such a painful exercise?” as reading through huge tables of output. He describes several options and points out the strengths and weaknesses of many creative visualization approaches.

“Using Predictive Modeling to Risk-Adjust Primary Care Panel Sizes” by Anders Larson: Anders describes the use of nontraditional techniques for traditional risk situations, specific to health insurance risk scores. He shows the advantage of an ensemble of smaller models utilizing gradient boosting machines as an alternative to the more traditional generalized linear model. Along the way he describes using cross-validation to avoid overfitting, and Anders concludes, “There is not a one-size-fits-all solution to risk adjustment.”

“Bayesian Inference in Machine Learning” by Denis Perevalov: Just because maximum likelihood estimations (MLEs) are fast and scalable does not mean they are the best choice in all machine learning situations. Denis takes us back to basics with Bayesian inference, which he shows may be a better choice for smaller amounts of data or data that are narrow in the longitudinal direction. Plus, he shows that Bayesian inference can make more precise predictions, and its confidence intervals of model parameters are more interpretable.

“Maximal Information Coefficient: An Introduction to Information Theory” by Bryon Robidoux: Before reading Bryon’s article, we might have assumed that a “bit” is a binary digit. This is unfortunate because a binary digit and a bit are different, and Bryon shows how in this summary of information theory. Along the way, he also introduces us to “nats” and “bans” and conditional entropy. Pattern matching is a lot more mathematical than we supposed, and we can draw upon information theory for a head start.

“Variable Selection in Predictive Modeling: Does It Really Matter?” by Kailan Shang: Are more variables always better when you are trying to build a predictive model? Kailan tells us no! In addition to the added complexity, you run into the presence of collinearity caused by too many variables, and this detracts from the robustness of your models. Kailan shows ways to reduce the number of variables and stresses the need to use human expert judgment at various stages of the process.

“The First SOA Annual Predictive Analytics Symposium—A Recommended Investment! (Whether or Not Your Employer Pays for It)” by Dave Snell: The section growth and the burgeoning interest in predictive analytics have reached a new milestone. We are planning to have our own annual special interest meeting, like the Health section, the Life and Annuity folks, and the Valuation actuaries. Dave describes this new SOA event and explains why you should sign up for it just about the time you get this newsletter. Don’t wait! Sign up now!

Enjoy the issue! ■



Dave Snell, ASA, MAAA is technology evangelist at SnellActuarialConsulting in Chesterfield, Mo. He can be reached at [dsnell@ActuariesAndTechnology.com](mailto:dsnell@ActuariesAndTechnology.com).



Kevin Jones, FSA, CERA, is associate actuary at Milliman in Buffalo Grove, Ill. He can be reached at [Kevin.Jones@Milliman.com](mailto:Kevin.Jones@Milliman.com).