

## Editorial

---

# SAS<sup>®</sup> Global Forum

Jennifer L. Waller<sup>a</sup> and Tyler C. Smith<sup>b</sup>

<sup>a</sup>*Augusta University, Augusta, GA, USA*

<sup>b</sup>*National University, San Diego, CA, USA*

Co-Editors of the MASA Special Issue on SAS Global Forum 2018.

Informatics solutions have enabled enormous amounts of data to be input, transported, integrated, and warehoused allowing for never-before seen opportunities for knowledge-based refinements in all disciplines including business, health, and finance (Herland et al., 2014; Chen et al., 2012). The explosion of available data has created a strong need for analytics and statistical adaptations catering to the growing size and complexity of data available (Haider, 2015; Donovan, 2008). This in turn has brought about a quickly evolving analytics space that seeks to catch up to the new data availabilities.

SAS<sup>®</sup> Global Forum has published white papers for decades that describe a continuum of analytic solutions for the unique to the well-used approaches being leveraged to analyze data using SAS. These papers have been written to provide a valuable, user-based foundation of knowledge that remain publicly available. In 2017, the Chair of SAS Global Forum Dr. Goutam Chakraborty submitted a proposal and requested the SAS Global Forum Executive Board approve a peer-reviewed journal supplement for the 2018 Conference. The Board approved this effort and partnership with the journal Model Assisted Statistics and Applications to publish a competed list of exceptional annual conference papers in a peer-reviewed setting.

There were 54 submitted papers to this peer-reviewed competition and 19 peer-reviewers recruited for this special issue of Model Assisted Statistics and Applications, 13 (4) 2018. Initial reviews by the guest editors reduced the number of potential papers for review to 24 based upon completeness, following guidelines for submission, and journal parameters. Seventeen authors indicated that a manuscript would be submitted for consideration, with 10 manuscripts submitted. Each of the 10 papers were then reviewed by 2–3 reviewers over a two-month period and final recommendations were submitted for each paper by each of the reviewers. The guest editors reviewed these suggested revisions which included reviewer indication of rejection or not. Revisions were sought by 8 of the papers and a period of 6 weeks was given for revisions to be completed. Once those secondary reviews were returned with again indication of rejection or not by the reviewers, authors were requested once again to complete revisions as necessary. This process was followed one additional time to determine the final 8 papers that were accepted to this journal supplement and we are pleased to present these in this issue of MASA. Topics include articles on: advanced machine learning with leveraging multi-source data for better prediction accuracy; a two-stage modeling approach utilizing gradient boosting and neural network data mining techniques; an in depth review of ridge regression and multicollinearity; disease classification using support vector machine modeling comparing to logistic regression; a sequential method of training a classifier for optimal utilization in real-world situations; resolving key identifier issues with fuzzy matching techniques to successfully merge, join and match less than perfect, or “messy” data; analyzing of customer feedback using best-practices of text mining and supervised sentiment analysis; and comparison of different models for forecasting emergency department visits with gradient boosting yielding best success.

As Guest Editors, we would like to thank the authors of all submissions but especially the authors of the final 8 papers, the reviewers, SAS Global Forum Executive Board, and the professionals at the Journal Model Assisted Statistics and Applications.

**References**

- Herland, M., Khoshgoftaar, T.M., & Wald, W. (2014). A review of data mining using big data in health informatics. *Journal of Big Data*, 1(2).
- Chen, H., Chiang, H.L., & Story, V.C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly* 36(4): 1165-88.
- Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2): 137-44.
- Donovan, S. (2008). Big data: teaching must evolve to keep up with advances. *Nature*, 455(1): 1.