David Jungreis

Pamplin Hall | Blacksburg, VA | Jungreis [at] VT [dot] EDU

EDUCATION

VIRGINIA TECH, Blacksburg, VA — August 2022-Current
Doctor of Philosophy: Finance

TEMPLE UNIVERSITY, Philadelphia, PA — May 2018
Master of Science: Statistics

UNIVERSITY OF PENNSYLVANIA, Philadelphia, PA — May 2014
Bachelor of Arts: Mathematics (Chemistry Minor)

PROFESSIONAL EXPERIENCE

J.F. TAYLOR, INC., Lexington Park, MD — October 2018-July 2022
Systems Analyst

- **Saved** a year of analysis time for the simulation sector of a two-trillion-dollar program by teaching the correct interpretations of hypothesis test p-values, effect sizes, statistical significance, and practical significance, refining the explanation to involve a well-received allusion to the *Princess and the Pea* fairy tale
- **Earned** funding on company overhead dollars to perform internal research and development on artificial intelligence, machine learning, and data science, leading investigations of nonparametric statistics, image processing, and regression modeling strategies that better the analysis of remote sensor output data—with software implementations written in both R and Python under *git* version control from the command line
- **Reduced** image feature space 87% via nonparametric, copula-generating comoment calculations that preserve mutual information and separability in t-distributed stochastic neighbor embedding (t-SNE) plots in order to maintain high supervised machine learning predictive performance, coding in R all feature extraction, data management, and Tensorflow/Keras convolutional neural network deep learning
- **Improved**, up to 50-fold, the power of hypothesis testing sensor fusion confusion matrix accuracy, applying multinomial logistic regression generalized linear models to expose performance against particular targets and ultimately basing the inferential statistics on a likelihood ratio "chunk test" I had to invent that used the Python packages *statsmodels* and *sklearn* for model fitting; *numpy* and *pandas* for data preparation and construction of a feature space; and *scipy* for calculating p-value probabilities—all using command line *git* for code version control
- **Sped** up the generation of an electromagnetic sensor data product delivery by extracting from partition means, variances, and sample sizes what those values would have been for all partitions combined, had those values been calculated when we had all of the data, thus saving the customer the weeks of time it would have taken to go back to an outside contractor with a request for further data analysis or a request for data to be shipped to my group
- **Eased** the process of recruiting my replacement as lead statistician by screening, interviewing, and giving feedback on candidate successors, including writing and asking technical questions (mixed between objective and open-ended) that examined for common fallacies about class imbalance, p-value interpretations, statistical model evaluation, and feature space covariate assumptions

HIGHLIGHTED PUBLICATIONS AND CITATIONS

D. Jungreis and S. Mukhopadhyay, "QDComparison: Modern Nonparametric Tools for Two-Sample Quantile and Distribution Comparisons." R package version 3.0. https://CRAN.R-project.org/package=QDComparison. 2019.

S. Lopez, et al., "Automated Identification of Abnormal Adult EEGs." IEEE Signal Processing in Medicine and Biology Symposium 2015. PMID: 27195311.

W. Abuzeid, et al., "Molecular disruption of RAD50 sensitizes human tumor cells to cisplatin-based chemotherapy." The Journal of Clinical Investigation 2009;119(7):1974-1985. PMID: 19487811.