Trenton Pulsipher

Sr Data Scientist
Henry Schein One, 1220 S 630 E, American Fork, UT 84003
tcpulsipher@hotmail.com 801-380-6878 trentonpulsipher

Employment

Henry Schein One

Sr Data Scientist (American Fork, UT) Apr 2017 - Present

As the lone data scientist at Henry Schein One I lead efforts to build predictive modeling via machine learning algorithms, promote data governance and business process documentation, and regularly share analytical best practices.

I create web-based tools to share statistical modeling visualizations with the sales, customer success, and marketing teams internally. Built in Linux these analytical tools allow me to share interactive JavaScript based apps via HTML pages in an R markdown/blogdown/Hugo generated website.

By applying automated machine learning predictive methods to customers at-risk of attrition we expect to enable savings of $10M+ within three years. I am implementing these same methods and framework to improve marketing messaging and sales lead generation. I also support the business operations team's effort to improve the data warehouse architecture, dimensional modeling, and reporting services (Tableau).

Category Partners

Sr Analyst (Idaho Falls, ID) Apr 2016 - Apr 2017

As the senior analyst of a five person startup I led the analysis, visualization/dashboard efforts of IRI/Nielsen syndicated data for grower/shipper clients in the produce category.

The Church of Jesus Christ of Latter-day Saints

Data Analyst / Statistician (Salt Lake City, UT) Dec 2014 - Apr 2016

As a data analyst in the Missionary Department I was responsible for all data science related analytical efforts within the department, including: data wrangling, modeling, predictive analytics, visualization, data discovery/exploration, and communication of results to the highest level of leaders/executives in the Church.

Pacific Northwest National Laboratory

Statistician (Richland, WA) Jul 2009 - Dec 2014

Implemented Bayesian model averaging (BMA) as an ensemble-based method to incorporate predictions from multiple experts or model systems. Successful application of BMA in modeling of response surfaces numerically generated from multiple expert-derived conceptual models characterizing sites for carbon sequestration, and have improved predictive performance for ensemble-based applications in image classification (computer vision), power grid demand modeling, pKa protein analysis, and sensor degradation signature discovery. Coauthored multiple

peer-reviewed journal articles and conference papers demonstrating expertise in BMA related applications.

Developed statistical methodology to spatially allocate weather similar regions using climate data. Reduction of the number of weather similar regions and of the number of building types will save climate modelers and power grid researchers at least hundreds of hours of computation and analysis time when using PNNL's complex regional earth systems simulation modeling framework, PRIMA. Implemented PAM clustering to adequately describe a subset of representative building types. Managed Big Data and parallelized statistical computations using Rhipe, an R package designed to interact with the Hadoop framework via Map/Reduce on PNNL's institutional supercomputer.

Performed multiple complex data analyses for other project work related to a variety of areas, including: building energy modeling, power grid modeling, bio-forensics of biowarfare growth media, river salmon population studies, air chemical analysis, uncertainty quantification for visualization of carbon sequestration, and several early drug development clinical trial related analyses.

Massachusetts General Hospital

Biostatistician (Boston, MA) Jun 2007 - Jul 2009

Tested models for early detection of ovarian cancer in clinical trial data. Examined longitudinal protein biomarker profiles from blood/serum. Calculated sensitivity and specificity for predicted data from logistic regression of potential biomarker combinations from cancer case/control studies.

Education

Brigham Young University, MS Statistics 2007

Brigham Young University, BS Statistics 2005

Publications

Journal Articles

Hejazi MI, Voisin N, Lui L, Bramer LM, Fortin DC, Hathaway JE, Huang M, Leung LR, Li HY, Patel PL, **Pulsipher TC**, Rice JS, Tesfa TK, Vernon CR, Zhou Y. 2015. "21st century United States emissions mitigation could increase water stress more than the climate change it is mitigating." *Proc Natl Acad Sci USA* 112(34):10635-40.

Gosink LJ, Bensema K, **Pulsipher TC**, Obermaier H, Henry MJ, Childs H, Joy K. 2013. "Characterizing and Visualizing Predictive Uncertainty in Numerical Ensembles Through Bayesian Model Averaging." *IEEE Transactions on Visualization and Computer Graphics* 19(12):2703-2712

Gosink LJ, Hogan EA, **Pulsipher TC**. 2013. "Bayesian Model Aggregation for Ensemble-Based Estimates of Protein pKa Values." *Proteins: Structure, Function, Bioinformatics* 82(3):354-363

Vlachopoulou M, Gosink LJ, **Pulsipher TC**, Ferryman T, Zhou N, Tong J. 2013. "An Ensemble Approach for Forecasting Net Interchange Schedule." *Power and Energy Society (PES) General Meeting 2013 IEEE*

Stephan CN, Amidan B, Trease H, Guyomarc'h P, **Pulsipher TC**, Byrd JE. 2013. "Morphometric Comparison of Clavicle Outlines from 3D Bone Scans and 2D Chest Radiographs: A Shortlisting Tool to Assist Radiographic Identification of Human Skeletons." *Journal of Forensic Sciences* 59(2):306-313

Nackos AN, Truong TV, **Pulsipher TC**, Kimball JA, Tolley HD, Robison RA, Bartholomew CH, Lee ML. 2011. "One-step conversion of dipicolinic acid to its dimethyl ester using monomethyl sulfate salts for GC-MS detection of bacterial endospores." *Anal. Methods* 2011(3):243-258

Yurkovetsky Z, Skates SJ, Lomakin A, Nolen B, **Pulsipher TC**, Modugno F, Marks J, Godwin A, Gorelik E, Jacobs I, Menon U, Lu K, Badgwell D, Bast Jr RC, Lokshin AE. 2010. "Development of a Multimarker Assay for Early Detection of Ovarian Cancer." *Journal of Clinical Oncology* 28(13):2159-2166

Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, Spiegelman CH, Zimmerman LJ, Ham AL, Keshishian H, Hall SC, Allen S, Blackman RK, Borchers CH, Buck C, Cardasis HL, Cusack MP, Dodder NG, Gibson BW, Held JM, Hiltke T, Jackson A, Johansen EB, Kinsinger CR, Li J, Mesri M, Neubert TA, Niles RK, **Pulsipher TC**, Ransohoff D, Rodriguez H, Rudnick PA, Smith D, Tabb DL, Tegeler TJ, Variyath AM, Vega-Montoto LJ, Wahlander A, Waldemarson S, Wang M, Whiteaker JR, Zhao L, Anderson NL, Fisher SJ, Liebler DC, Paulovich AG, Regnier FE, Tempst, Carr SA. 2009. "Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma." *Nature Biotechnology* 27:633-641

Other Publications

Baker NA, Dowling C, Gosink L, **Pulsipher TC**, Sansone SA. 2014. "Informatics Approaches to Data Preservation and Analysis in Protein Electrostatics." *Biophysical* 108-2(369A)

Awards

R&D 100 Award

2015 R&D Magazine. "Power Model Integrator: A system for more accurate energy forecasts."

Patent

"System and Method of Designing Models in a Feedback Loop." Patent issuer and numberus 13/869,290