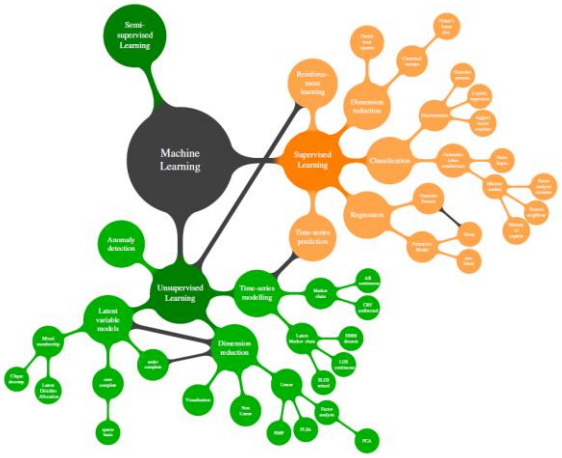
WCU WEST CHESTER UNIVERSITY

Applied Statistics Newsletter

Fall 2018

Industry Insight

by Amos Odeleye, Alumus 2011



What is Machine Learning?

Simply speaking, I will define Machine Learning as any approach- equation, method or algorithm, that is used to "learn from data".

These approaches encompass three scientific disciplines namely: Statistics, Mathematics, and Computer Science. Other terminologies used synonymously are Artificial Intelligence (AI), Data Mining, or Data Science.

- *A Statistician may define Machine Learning as a set of methodologies for modeling and analyzing complex, large, or "big" datasets. Hence, Multivariate Statistical Analysis, or "Statistical Learning."*
- *A Mathematician may describe Machine Learning as set of formulated equations from given phenomenon to explaining data or used in forecasting. Hence, Mathematical Modeling*
- *A Computer Scientist may express Machine Learning as a set of algorithms used in understanding and prediction of data. The term grew out of Computer Science.*

As Statisticians, to provide solutions to solve real-life problems, we are not limited to statistically-derived methods, so explore and let your *machines learn from your data*.

Further reading:

- *de Mello,Rodrigo Fernandes, Ponti, and Moacir Antonelli.* ***Machine Learning-A Practical Approach on the Statistical Learning Theory***
- *Trevor Hastie, Robert Tibshirani, and Martin Wainwright.* ***Statistical Learning with Sparsity-The Lasso and Generalizations***
- *Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani.* ***An Introduction to Statistical Learning***
- *Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein.* ***Introduction to Algorithms***
- *Trevor Hastie, Robert Tibshirani, Jerome Friedman.* ***Elements of Statistical Learning***
- *Simon Rogers and Mark Girolami.* ***A First Course in Machine Learning***
- *Christopher M. Bishop.* ***Pattern Recognition and Machine Learning***
- *Kevin P. Murphy.* ***Machine Learning-A Probabilistic Perspective***
- *David Barber.* ***Bayesian Reasoning and Machine Learning***
- *Yoshua Bengio, and Aaron Courville.* ***Deep Learning***

Alumni Updates

Mike Borish

Graduation year: 2007
Contact: mikeborish@gmail.com

Current Position and Job Description
Senior Product Strategy Analyst at TD Bank (April 2017 to Present)
Prepare and present Credit Risk and Marketing Campaign Strategy Analytics in support of the Nordstrom Credit Card Partnership.

Previous Position and Job Description
Senior Manager of Audit Sampling and Analytics (2007-2017)
Consulted state government agencies on Audit Sampling and Estimation Standards for over 200 Unclaimed Property Audits of Fortune 1000 companies, resulting in millions of dollars being reunited with their rightful owners

Internship
Statistical Programmer at Omnicare Clinical Research (2007)

Marshal Ma

Zhen-qiang (Marshal) Ma is a proud graduate of West Chester University's Applied Statistics program (2008). His decision to pursue a Master of Science in Applied Statistics began while he was working as an epidemiologist at Montgomery County Health Department in Pennsylvania. During his tenure at the health department, Marshal identified that much of the data collected was not analyzed using the appropriate techniques. The applied statistics graduate program offered him an opportunity to learn and apply his education while working full time.

After completing his master's degree, Marshal initiated and created the first systematic Montgomery County Health Department Annual Report. He also applied statistical methods in cancer cluster investigations, disease outbreak investigations, and the National Children's Study. After Montgomery County Health department, Marshal worked at SDI health (now a part of IQVIA) as a business consultant/statistical analyst. Marshal is currently a division director at the Pennsylvania Department of Health where he oversees chronic disease related issues. He is also an adjunct professor of epidemiology at the University of Pittsburgh Graduate School of Public Health. At the Pennsylvania DOH, Marshal created the first state chronic disease report using multiple data sources and has published many journal articles, including a collaborating with

his mentors at the West Chester University statistics program. He is currently collaborating with academic researchers in the NIH, CDC, and ATSDR on many funded projects and serves on various state and national committees.

Marshal credits the applied statistics graduate program with giving him the knowledge and foundation to be successful in his career.

Kira Eileen Pugliese

Kira is a fiercely proud graduate of West Chester University's Applied Statistics program (2014). Originally a biochemist, Kira realized she much preferred quantitative analysis of investigational data so she left the bench to pursue a master's degree in Applied Statistics. While at West Chester University, Kira served as a Graduate Assistant for the department while completing a co-op with AstraZeneca as a Quantitative Commercial Insight analyst for its Cornerstone brands.

Though she worked on AstraZeneca's commercial side which was a new world to her, she realized that she could leverage common concepts from her scientific and statistical training to help her team maximize brand performance. The co-op led to a full-time job and Kira's responsibilities continued to grow as the organization accelerated its analytical capabilities. When a robust Advanced Analytics team was built, Kira led advanced analytics projects for several products in oncology including creation of the first broad-scale marketing mix model. Due to the challenges inherent to health-care data in oncology (such as very small patient populations, individualized treatment algorithms, blocking and missing data), Kira was unable to adopt traditional marketing-mix modeling methodology. Instead, she applied a method that is used in health economics to evaluate cost-effectiveness of healthcare strategies as patients transition from health to illness to death. With this statistical framework she was able to help her teams understand which marketing tactics were most impactful in moving a healthcare provider from non-user to trialist to repeater. Optimization of marketing strategies ultimately helped TAGRISSO reach global blockbuster status just two years after its first approval.

Kira has since expanded upon her breadth of experience in AstraZeneca's Insight and Analytics function by serving as a commercial forecast leader for both oncology and primary-care products. She introduced time-series models to the oncology forecasting team and is currently working cross- functionally to integrate Monte Carlo simulation into forecasting best practices. In an effort to boost her business acumen, Kira has recently started an MBA program at West Chester University which is helping her elevate the strategic value of her contributions to the organization.

Kira is grateful for all of the opportunities that WCU Applied Statistics has presented to her including the privilege of serving as a Graduate Assistant which she warmly reflects back on as a highlight of her time in the program. Among many other things, she credits West Chester University Applied Statistics with fostering her love of numbers and imparting the skills required to be a successful statistician.

Honoring Our Current Students

Sunday November 4th the Department of Mathematics hosted their annual Awards Banquet. This year we honored the academic excellence of the Applied Statistics graduate class of 2017. This remarkable group of students has broken a number of records. First off, they are our largest class thus far, with well over 60 students. Secondly, they have received a record number of internships. Placements included Cobbs Creek Healthcare, West Chester University, The University of Pennsylvania, Lawyermetrix, Stroud Water Research, Chester County Hospital, Astra Zeneca, Endo Pharmaceuticals, Citicard and FMD K&L. Finally, we had a record number of students honored as recipients for their academic excellence at the Awards Banquet The six recipients this year were Jialai Chen, Dana Heleniak, Patrick Murphy, Michael Pol, Joe Sobieski and Yuping Zhang.

Get Involved!

Want to volunteer with the Applied Statistics program or Applied Statistics Alumni Group? Consider helping out with the newsletter, planning our next networking event, offering an internship for current students, or organizing an educational session for students, alumni and faculty.

If you have any question, comments or suggestions, please contact:

- Sopheara Peoples sopheara.siek@gmail.com
- John Bragger john_bragger@comcast.com
- Amos Odeleye statisticalliance@gmail.com

Continue your education with our upcoming courses

Spring:

STA 533 -- Longitudinal Analysis -- Introduction to the application and theory of models for clustered and longitudinal data. Course will address the analysis for both continuous and categorical response data. Course will be held in the statistics lab and use the statistical software package SAS. Other software such as R, HLM, SPSS, MIXORMIXREG may be introduced.

STA 544 -- Taught by two experienced leaders in the field (Professors Sese Abhulimen and Brian Wynne) from nearby internship partner companies, this course will provide practical hands on knowledge and experience on the application of statistical methods and techniques to Big Data. It will prepare students with a good understanding of the workins of business analytics and applications within retail, pharmaceutical, and other industries.

Summer One:

STA 532 -- Survival Analysis -- This course will provide students with the knowledge and tools to conduct a complete statistical analysis of time to event data. Students will get experience using common methods for survival analysis, including Kaplan-Meier Methods, Life Table Analysis, parametric regression methods, and Cox proportional Hazard Regression. Additional topics include discrete time data, competing risks, and sensitivity analysis.

Summer Two:

STA 542 -- Statistical Methods for Observation Studies -- In the assessment of the association between a predictor and a response confounding by another factor might yield wrong answers. One standard technique to protect against confounding is randomization, which is the standard for conducting randomized clinical trials (RCT). In the setting where randomization cannot be applied, such as cohort or case-control studies, the potential for confounding exists; therefore, analytical techniques must be developed to address this potential confounding. These studies where the respective predictor is observed (i.e. gender, case versus control, etc...) rather than randomized (i.e. drug versus placebo, Treatment 1 versus Treatment 2, etc...) are referred to as observational studies. This course will cover statistical methods for the design and analysis of observational studies. Students will be exposed to discussion of differences between experimental, observational, and quasi-experimental studies. Techniques to assess statistical effects while addressing confounding (both measured and unmeasured) and selection bias will be introduced. Various techniques introduced are: propensity scores, inverse probability weighting, instrumental variables, Marginal Structural Models, Structural Nested Mean Models. Students additionally will be introduced to the Rubin Causal Model framework in the assessment of Causal effects.

\

Applied Statistics Newsletter