Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie or Die

Seewon Ryu, PhD

Department of Health Policy and Healthcare Management, Inje Institute of Advanced Studies, Seoul, Korea
seewon@inje.ac.kr

Author: Eric Siegel

Year: 2013

Publisher: John Wiley & Sons, Inc.

ISBN: 978-1-118-35685-2

Hospital captures daily large amounts of data about their customers or patients, suppliers, and operations. Health insurance organizations also have large claims data which called as big data—large pools of data that can be captured, communicated, aggregated, stored, and analyzed [1]. We can analyze big health data to detect signals that is useful for patients and healthcare service management, although the data has quality problems. It is increasingly the case that healthcare innovation and growth could take place with predictions analyzing big data [1].

I introduce the book by referring book description and author's interview with amazon.com.

The author makes the how and why of predictive analytics (aka PA) understandable and captivating in various fields of big data through the book. Although it is not targeted only for researcher or stakeholders in healthcare area, the book would be useful for us because it is targeted from the small to large business owner, entrepreneurs, other PAers and us common folk who want to further understand how computerized data research is analyzed to predict specified outcomes and scenarios.

The book breaks down predictive analytics into seven chapters. Cause and effect charts, illustrations along with a few comics and a glossy centerfold divulge cases of predictions in advertising, finance, healthcare, fraud, insurance, government, employment and personal venues. Some topics discussed explain ways to increase consumer buying, limit bank loan defaulting or paying off, anticipate employees quitting or clients dropping cellphone coverage along with collecting online blogs, social networking and risk information. Each chapter includes sections of “what’s predicted” and “what’s done about it” to show the correlation of PA and gathered

data.

The author explains the art of predicting has five effects that include:

- 1) A little prediction goes a long way,
- 2) Data is always predictive,
- 3) Induction is reasoning from detailed facts to general principles,
- 4) Ensembles compensate for limitations, and
- 5) Persuasion can be predictable through outcomes.

Using the predictive models of large corporations such as Target, Hewlett-Packard, Chase Bank, Netflix and Telenor along with John Elder's stock market techniques, Jeopardy!'s Watson computer, Kaggle's competitions, and Obama's second term presidential campaign, we can learn the ins and outs of predicting through collecting and interpreting simple to complex data.

By entrusting computers to make decisions, privacy concerns are bought up, prejudices are determined and effects are manipulated when machine learning becomes the translated voice of data. Artificial intelligence can often limit overlearning, crowdsourcing and correlation pitfalls, but will it be able to always correctly interpret language, emotions and feelings of humans as it influences, persuades and molds us?

With even the book's title been subjected to analysis and written sometimes humorously of the writer's own experience of stolen identity and mockery of his geekness, it is an excellent source to any reader that sees computers overtaking and controlling our every move as we continue to be co-dependent on them as we happily benefit from increased information and understanding, attain higher profits and enjoy an easier lifestyle through such a conglomerate of PA data bytes.

In the book, Siegel explains real-world examples on how organizations are turning big data into meaningful metrics. You have been predicted—by companies, governments, law-enforcement, hospitals and universities. Their computers say, "I knew you were going to do that!" These institutions are seizing upon the power to predict whether you're going to click, buy, lie, or die.

Predicting human behavior, by analyzing accumulated health service data not so interested until now, would fortify healthcare. Accumulated in large part as the by-product of routine tasks, data is the unsalted, flavorless residue deposited en masse as organizations churn away. Surprise! This heap of refuse is a gold mine. Big data embodies an extraordinary wealth of experience from which to learn.

Siegel assert that predictive analytics is the science that unleashes the power of data. With this technology, the comput-

er literally learns from data how to predict the future behavior of individuals. Also, he describes that perfect prediction is not possible, but even lousy predictions can be extremely valuable.

In the book, Eric Siegel reveals the power and perils of prediction:

- What unique form of mortgage risk Chase Bank predicted before the recession.
- Predicting which people will drop out of school, cancel a subscription or get divorced before they are even aware of it themselves.
- Why early retirement decreases life expectancy and vegetarians miss fewer flights.
- Five reasons organizations predict death, including one health insurance company.
- The way United States Bank and European wireless carrier Telenor calculate how to most strongly influence each customer.
- How companies ascertain untold, private truths—how Target figures out you're pregnant and Hewlett-Packard deduces you're about to quit your job.
- How judges and parole boards rely on crime-predicting computers to decide who stays in prison and who goes free.
- What's predicted by the BBC, Citibank, ConEd, Facebook, Ford, Google, IBM, the IRS, Match.com, MTV, Netflix, Pandora, PayPal, Pfizer, and Wikipedia.

Siegel insist that predictive analytics as a truly omnipresent science affects everyone, every day. Although largely unseen, it drives millions of decisions, determining who to call, mail, investigate, incarcerate, set up on a date, or medicate.

Predictive analytics transcends human perception. The final chapter answers the riddle: What often happens to you that cannot be witnessed, and that you can't even be sure has happened afterward—but that can be predicted in advance? Whether you are a consumer of it—or consumed by it—get a handle on the power of Predictive Analytics.

Have you ever heard that those who buy diapers are more likely to buy beer too? How about that when stapler is sold at a retailer, it is a good indicator that a company has hired a new employee? Would you be surprised to learn that those using their credit card to buy a drink in a "drinking establishment" are more likely to miss their credit card payment?

A few other surprising facts readers will find in the book include:

- Clinical researchers predict divorce with 90 percent accuracy.
- Researchers employ machine learning to predict Hollywood blockbusters and hit songs.

- Life insurance companies predict age of death to decide whether to approve and then price a policy application.
- The state of Maryland uses predictive models to detect inmates more at risk to be perpetrators or victims of murder.
- University of Phoenix predicts which students are at risk of failing a course and then target them with intervention measures.

To fully describe what predictive analytics are and how such data is used, Siegel uses these facts and case studies such as big box retailer, Target, predicting which female customers will have a baby in coming months so that they can market relevant items to the expectant parents. When you go to the grocery store or drug store and you get coupons printed that seem to offer just the products you may be interested in, that is another every day use of PA that the average consumer may encounter.

Siegel writes, “I was at Walgreens a few years ago, and upon checkout, an attractive, colorful coupon spit out of the machine. The product it hawked, pictured for all of my fellow shoppers to see, had the potential to mortify. It was a coupon for Beano, a medication for flatulence.” The author had developed lactose intolerance, and the drug store was recommending products that might be compatible with his medical condition. He describes PA as benefiting both the consumer and the organization by empowering it “with an entirely new form of competitive armament.”

Siegel spends a lot of the book going through the different predictive models and how the data is extracted and used. This type of data mining is being used in numerous ways, not just for consumerism. He provides examples from family and life, healthcare, crime fighting and fraud detection, staff and employees and human language understanding. While the book answers so many questions about how marketers know so much about consumers, it's not just a book for marketers. It is one of those business books that cross many departments including sales, research and development, hu-

man resources, customer service among others.

There are many big data in healthcare area such as health insurance corporations, hospitals, bioinformatics research institutions, and disease management and control institutions. PA would contribute to intervene in utilization of healthcare service, guide new way of supply of healthcare service, enhance lifestyle, and prolong life expectancy. The book introduces 147 examples of predictive analysis including healthcare industry. These mini-case studies easily will inspire and broaden understanding and utilization of predictive analytics to health professionals and scholars interested with big data analysis.

Clinical data has issues on ownership of medical records and also legal issues related to data access, pooling, and use [2]. These issues may be hurdle to apply predictive science with pooled big data, but we can utilize predictive analytics effectively to the limited data.

Practitioners in hospital can use the book as a guide to invent new way of service and business by using the amount of refuse data. Researchers also may acquire insight on research directions to predictive analysis. Big data aligned from various sources will give meaningful implications to policy maker and practitioners about one's health behavior as the case diaper.

References

1. Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, et al. Big data: the next frontier for innovation, competition, and productivity. New York (NY): McKinsey Global Institute; 2011.
2. Goodby AW, Olsen L, McGinnis M; Institute of Medicine (US) Roundtable on Value & Science-Driven Health Care. Clinical data as the basic staple of health learning: creating and protecting a public good. Washington (DC): National Academies Press; 2010.