

# Predictive Modeling in Practice

Tracey De Poalo  
Sprint  
Kansas City, MO

Jeremy Howard  
Khosla Ventures  
Menlo Park

## ABSTRACT

It's not often we get to peer into the details of how big corporates use predictive modeling in practice. In this talk, Sprint's Head of Predictive Modeling, Tracey De Poalo, will talk about the process she developed using SAS and logistic regression to build a wide range of models. Jeremy Howard will discuss his experience as a consultant at Sprint comparing R and random forests to the existing process, and will show the pros and cons of each approach.

The talk will also cover the "real world" issues that Tracey has dealt with in creating a reusable process, including data mart development, sampling, testing, reporting, and implementation with internal customers.

Sprint's existing process is the best that Jeremy has seen in industry, and shows a lot of best practices in data structure, documentation, testing, automation, and customer interaction, and modeling.

## Categories and Subject Descriptors

H.4.0 Information Systems, INFORMATION SYSTEMS APPLICATIONS: General.

## Keywords

Predictive modeling; machine learning; logistic regression; random forests; real-world applications

## Bio

Tracey has been managing the Predictive Analytics team at Sprint for six years. Her educational background includes a Masters in Statistics from Florida State University and a Masters in Industrial Design from Georgia Institute of Technology. A deep understanding of design methodology allows Tracey to bring creative solutions to complex analytical problems. She has introduced multiple, robust procedures that have made predictive analytics at Sprint more efficient and productive. Tracey also has extensive analytical experience in the credit industry working for Equifax and Merkle.

Jeremy Howard is an entrepreneur, business strategist, developer, and educator. He is the founder of Strategic Data Science, which aims to make machine learning accessible to non-experts, and is a Distinguished Research Scientist at the University of San Francisco. He is also a faculty member at Singularity University, and a Young Global Leader with the World Economic Forum. Previously he was the President and Chief Scientist of the data science platform Kaggle. He was the founding CEO of two successful Australian startups (FastMail, and Optimal Decisions Group). Before that, he spent 8 years in management consulting, at McKinsey & Co, and AT Kearney. Jeremy has invested in, mentored, and advised many startups, and contributed to many open source projects.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).

*KDD'14*, August 24–27, 2014, New York, New York, USA.  
ACM 978-1-4503-2956-9/14/08.  
<http://dx.doi.org/10.1145/2623330.2630821>