EPFL

Data Science for Managers

DS fM

The Organizational Value of Training Managers in Data Science Methods

Kenneth Younge
Associate Professor, EPFL

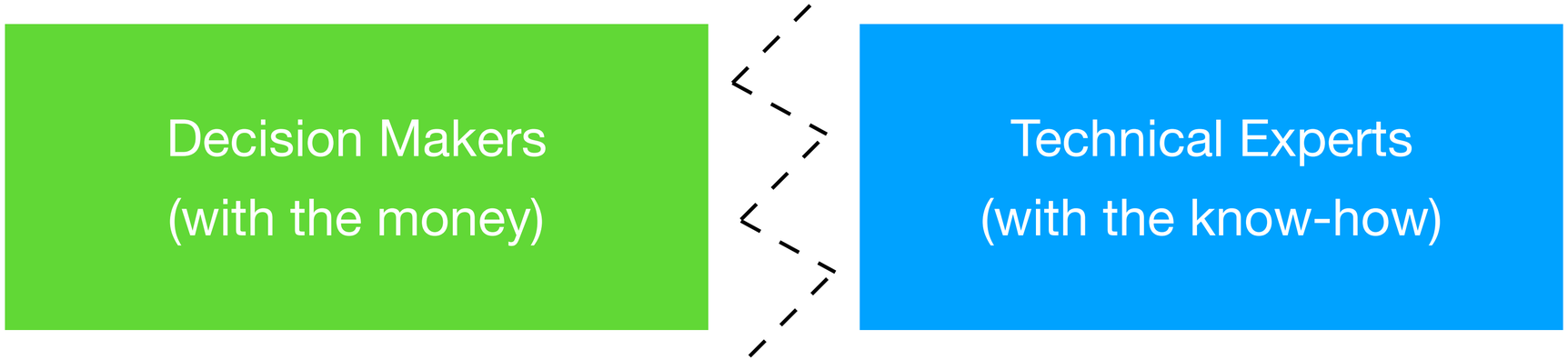
As an economist, I have a shocking discovery to share with you…

“Managers decide which ideas are actually allocated resources.”

It is the manage matters the most for the overall new projects.

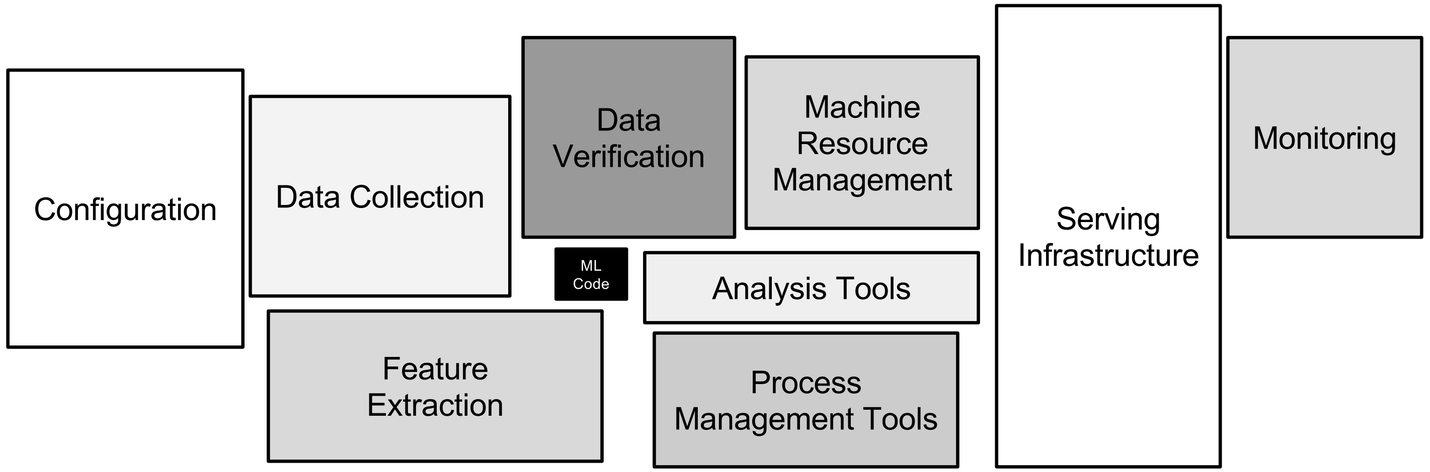
(Mollick, 2012)

What does this have to do with Data Science?



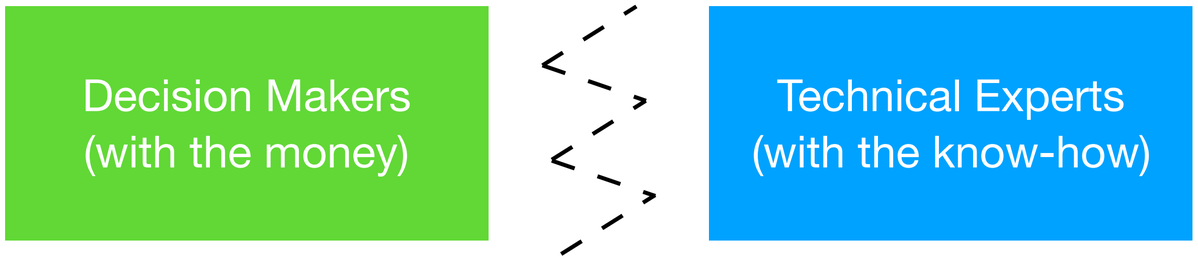
The two sides often view the “AI” opportunity quite differently.

Communication and coordination is difficult when the ML environment is very complex.



Sculley, David, et al. "Hidden technical debt in machine learning systems." *Advances in neural information processing systems.* 2015. (Google)

Is your firm "*Data Science Ready*?"



Who prioritizes objectives and assesses risks?

Who selects a vendor or a platform?

Who evaluates results and maintains solutions going forward?

Who recruits new employees?

How do you become *Data Science Ready*?

Firms need to upskill a broader range of their executives, managers, engineers, professionals, and domain experts so they share the same basic concepts and vocabulary.

There are many paths to this goal….

DS
fM

For some of you, it’s time to go back to school !

EPFL offers two continuing education courses for onsite Data Science (ML) instruction:

2-Day Executive Fast Track

~~February 6 - 7, 2020~~ - Full
April 2 - 3, 2020
June 25 - 26, 2020
October 8 - 9, 2020

5-Day Technical Boot Camp

March 2 - 6, 2020 (Filling fast...)
June 8 - 12, 2020
September 7 - 11, 2020

Why take a class?

In person?

At EPFL?

DS
fM


Data Science for Business
Provost & Fawcett
O'REILLY
James · Witten · Hastie · Tibshirani
An Introduction to Statistical Learning

Prof. Kenneth Younge is an Associate Professor at EPFL, the Chair of the Technology and Innovation Strategy Lab, and Program Director for DSFM. He has started four companies and worked in the roles of Director of Development, Consultant, CTO, and President.

Professor Younge currently teaches the Masters course on **Data Science for Business**, the Doctoral course on **Computational Methods for Management**, the IML course on **Data Science for Logistics**, and an eMBA course on **Technology and Innovation Strategy**. His research focuses on computational economics and digital transformation. His doctoral students and post-doctoral researchers collaborate with a wide range of Swiss and US companies on ongoing research projects.

We organize **interactive lunch sessions** during the DSFM Boot Camp. Additional professors and industry experts join the class to meet with small groups of DSFM participants to discuss particular topics of interest. Each discussion leader is an expert in a given area. Examples of discussion leaders include:

Prof. David Atienza is an Associate Professor at EPFL and expert on embedded systems for the Internet of Things (IoT). He leads a discussion on how smart wearables, wireless sensors, edge computing, and embedded machine learning work together to create new business opportunities.

Prof. Dimitrios Kyritsis is an Adjunct Professor at EPFL and Director of the Doctoral Program on Robotics, Control and Intelligent Systems. He is an expert on the management of data and information flows, and D-I-K (Data-Information-Knowledge) transformations all along the lifecycle of products.

Prof. Bob West is an Assistant Professor at EPFL and head of the Data Science Lab in the School of Computer and Communication Sciences. His research aims to make sense of Big Data collected from the Web, such as server logs, social media, wikis, online news, online games, etc.

Prof. Alex Biedermann is an Associate Professor at the UNIL Ecole des Science Criminelles and an expert on decision-making under uncertainty. He leads a discussion on how computational methods can support a more systematic approach for automating decisions.

Prof. Chris Tucci is the former Dean of the College of Management at EPFL and an expert on issues of design thinking and digital transformation. He leads a discussion on how firms construct data-driven strategies to transition to new technologies, business models, and organizational forms.

Prof. Negar Kiyavash is a Full Professor at EPFL and Chair of Business Analytics in the the College of Management. She is an expert on causal inference from networked big data and teaches on advanced topics in machine learning, artificial intelligence, optimization, and data science.

Prof. Jeffrey Kuhn is an Assistant Professor at the University of Carolina and a US patent attorney. He leads a discussion on how AI algorithms and proprietary methods can best be protected as either intellectual property or trade secrets.

Dr. Christopher Bruffaerts is a lecturer at the College of Management at EPFL and instructor for the Masters course on Data Science in Practice. He has worked on customer analytics, fraud detection, and big data technologies at BNP Paribas Fortis, Credit Suisse, and UPC.

These are “hard” and “challenging” courses !

2-Day Fast Track

Day 1:

Core ML Concepts
Core ML Methods
Neural Networks
Cloud Computing
Data Engineering

Day 2:

AI Solutions
AI Strategy
Digital Transformation

5-Day Boot Camp

Foundational concepts:

- Data sampling, measurement, and wrangling
- Exploratory data analysis
- Data description, visualization, and graphing
- Bias, variance, and the bias-variance tradeoff
- Model validation and model cross-validation
- Hyperparameter tuning and information leakage
- Model evaluation and comparison
- Model weighting of costs and benefits
- Ensemble learning and meta-learning
- Predictive labeling and data augmentation
- Data-driven business models
- Big Data, Map-Reduce, and Spark
- Virtual Machines and Cloud Computing
- Strategic Planning for a Digital Transformation
- The management of talent and strategic Human Capital

Methods and models:

- Normalizing and standardizing data
- Linear and Log-Linear models
- Non-parametric models, splines and locally-linear models
- Nearest neighbor and similarity models
- Agglomerative clustering and K-means clustering
- Decision trees, bagging, boosting, and random forests
- Dimension reduction, PCA, t-SNE, and manifold projections
- Support Vector Machines
- Text as Data and Natural Language Processing (NLP)
- Word Embeddings and Latent Topic Modeling
- Feed-Forward Neural Networks
- Convolutional Neural Networks
- Recurrent Neural Networks, LSTMs, Bi-Lateral LSTMs
- Generative Adversarial Networks
- Reinforcement Learning

DS
fM

DS
fM

dsfm.ch