The AI Playbook
Mastering the Rare Art of Machine Learning Deployment

By Eric Siegel

Notes for CHAPTER 0

These notes include references, plus resources for further learning. For all the chapters' notes as well as information about the book in general, access the book's website at www.bizML.com.

On ML combating wildfires:
https://www.thestar.com.my/tech/tech-news/2021/11/14/can-artificial-intelligence-help-improve-wildfire-recovery
https://www.weforum.org/reports/the-next-frontier-in-fighting-wildfires-fireaid-pilot-and-scaling
https://www.forbes.com/sites/jamesconca/2021/08/31/capturing-wildfires-from-space--pnnls-radrfire/
https://www.scientificamerican.com/article/ai-could-spot-wildfires-faster-than-humans/
https://www.wsj.com/articles/california-firefighters-tap-ai-for-an-edge-in-battling-wildfires-11601544600
https://www.voanews.com/a/new-artificial-intelligence-solutions-developed-to-combat-wildfires/7282474.html

On ML combating climate change:
https://predictiveanalyticsworldclimate.com/
https://www.aitheology.com/2021/02/01/how-scientists-are-using-ai-to-protect-the-amazon/

Other risk mitigation with ML:
https://www.nytimes.com/2022/04/19/technology/ai-road-car-safety.html
https://www.acusensus.com/case-studies/

On the high potential gains of retention when there is a high turnover:
Predictive Analytics' Killer App: Retaining New Customers:
https://www.predictiveanalyticsworld.com/customer-retention/

The Automation Paradox: "When computers start doing the work of people, the need for people often increases."
https://www.theatlantic.com/business/archive/2016/01/automation-paradox/424437/

The automation paradox: "As systems become more automated, humans lose some of their skill with the system, resulting in more automation."
https://www.youtube.com/watch?v=oMTb7u93mSI

Another "automation paradox": “By taking away the easy parts of [the] task, automation can make the difficult parts of the human operator’s task more difficult.... setting them straight when they go wrong requires deeper and deeper expertise.”
https://www.forrester.com/blogs/beware-the-automation-paradox/

On the origin of "the automation paradox": British cognitive psychologist Lisanne Bainbridge, in an influential paper of 1983, was thinking of machine control in process industries when she observed ‘the irony that the more advanced a control system is, so the more crucial may be the contribution of the human operator.’
https://ieeexplore.ieee.org/document/8013079
https://www.sciencedirect.com/science/article/abs/pii/0005109883900468?via%3Dihub
https://web.archive.org/web/20200717054958if_/https://www.ise.ncsu.edu/wp-content/uploads/2017/02/Bainbridge_1983_Automatica.pdf
https://personalmba.com/paradox-of-automation/#:~:text=The%20Paradox%20of%20Automation%20says,it's%20fixed%20or%20shut%20down.

Telenor example of churn modeling:

Suresh Vittal, “Optimal Targeting through Uplift Modeling: Generating Higher Demand and Increasing Customer Retention While Reducing Marketing Costs,” Forrester Research White Paper, 2008.
http://www.crmxchange.com/uploadedFiles/White_Papers/PDF/Optimal_Targeting_with_Uplift_Modeling_white_paper.pdf

Suresh Vittal, “Optimizing Customer Retention Programs,”
Forrester Research White Paper, 2008.
https://www.forrester.com/report/optimizing-customer-retention-programs/RES44400?objectid=RES44400
https://docplayer.net/2015695-For-direct-marketing-professionals.html

“The problem is that most executives are selected for their ability to talk to other people… Those who develop ML solutions to business problems are selected for their ability to talk to machines.”
–Mihnea Moldoveanu, Management Professor, University of Toronto
Why AI Underperforms and What Companies Can Do About It:
https://hbr.org/2019/03/why-ai-underperforms-and-what-companies-can-do-about-it

NOTE: The author has participated in a more recent project surveying data scientists on model deployment success rates subsequent to the project for which results are reported in Chapter 0. Here it is:

> **Survey: Machine Learning Projects Still Routinely Fail to Deploy**
> https://www.kdnuggets.com/survey-machine-learning-projects-still-routinely-fail-to-deploy

Survey reported in the book -- conducted by the author and KDnuggets -- Models Are Rarely Deployed: An Industry-wide Failure in Machine Learning Leadership
https://www.kdnuggets.com/2022/01/models-rarely-deployed-industrywide-failure-machine-learning-leadership.html

"Average ROI on enterprise-wide [AI] initiatives is just 5.9%,—well below the typical 10% cost of capital."
https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/ai-capabilities

Deployment as a Critical Business Data Science Discipline
https://hdsr.mitpress.mit.edu/pub/2fu65ujf/release/2

Rexer Analytics' survey of data scientists
https://www.rexeranalytics.com/data-science-survey

Research from MIT Sloan Management and Boston Consulting Group
https://sloanreview.mit.edu/projects/expanding-ais-impact-with-organizational-learning/

Earlier research from MIT Sloan Management: "Many AI initiatives fail."
https://sloanreview.mit.edu/projects/winning-with-ai/

Gartner estimate - from "85% of big data projects fail, but your developers can help yours succeed"
https://www.techrepublic.com/article/85-of-big-data-projects-fail-but-your-developers-can-help-yours-succeed/

According to a survey of senior executives by ESI ThoughtLab, "the average return on all AI investments by company is only 1.3%" and "only 20% of AI projects are in widespread deployment."
From: Survey of senior executives by ESI ThoughtLab
https://econsultsolutions.com/wp-content/uploads/2020/09/ESITL_Driving-ROI-through-AI_FINAL_September-2020.pdf

McKinsey’s AI Index reveals that the gap between leaders and laggards in successful AI and analytics adoption, within as well as among industry sectors, is growing
https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2020
https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/how-artificial-intelligence-can-deliver-real-value-to-companies
https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact

Only 21% of business and IT professionals said their AI initiatives were in production (Gartner), although another Gartner result showed 53% of AI projects making it from prototypes to preduction (albeit, the definition of AI includes chatbots and the like).
https://www.gartner.com/en/newsroom/press-releases/2020-10-01-gartner-survey-revels-66-percent-of-orgnizations-increased-or-did-not-change-ai-investments-since-the-onset-of-covid-19
https://www.gartner.com/en/newsroom/press-releases/2020-10-19-gartner-identifies-the-top-strategic-technology-trends-for-2021

McKinsey: "out of the 3,073 respondents, only 20 percent said they had adopted one or more AI-related technology at scale or in a core part of their business."
https://www.mckinsey.com/~/media/mckinsey/industries/advanced%20electronics/our%20insights/how%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/mgi-artificial-intelligence-discussion-paper.ashx

Why do 87% of data science projects never make it into production?
https://venturebeat.com/2019/07/19/why-do-87-of-data-science-projects-never-make-it-into-production/

A recent McKinsey Global Survey, for example, found that only about 15 percent of respondents have successfully scaled automation across multiple parts of the business. And only 36 percent of respondents said that ML algorithms had been deployed beyond the pilot stage.
https://www.mckinsey.com/business-functions/operations/our-insights/operationalizing-machine-learning-in-processes

Global AI Survey: AI proves its worth, but few scale impact
https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact

In 2020, there was actually “no increase in AI adoption,” according to McKinsey’s The State of AI.
https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2020

Why 90 percent of all machine learning models never make it into production https://towardsdatascience.com/why-90-percent-of-all-machine-learning-models-never-make-it-into-production-ce7e250d5a4a

"Only 22 percent of companies using machine learning have successfully deployed a model, the study found."
https://info.deeplearning.ai/the-batch-companies-slipping-on-ai-goals-self-training-for-better-vision-muppets-and-models-china-vs-us-only-the-best-examples-proliferating-patents

McKinsey: As little as 10% of potential value has been captured by analytics in certain sectors.
https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world

Only 29.2% of companies report achieving transformational business outcomes from AI and data initiatives.
https://hbr.org/2021/02/why-is-it-so-hard-to-become-a-data-driven-company

68% of ML practitioners admitted abandoning 40-80% of their experiments in the past year due to mismanagement.
https://www.comet.com/site/ty/report-2021-machine-learning-practitioner-survey/

13 percent of data scientists succeed in deploying their models, a total that hasn't improved much over the last eight years.
https://www.datanami.com/2017/09/12/alteryx-tools-aims-speed-model-deployment/

42% of data scientists affirmed that "Data science results not used by business decision makers":
The impact of increased digitization on the data science field
https://www.sas.com/content/sascom/en_us/offers/22q1/accelerated-digital-transformation-datascientist-ebook.html

"The greatest problems with artificial intelligence are not primarily technical, but rather how to achieve value from the technology."
https://www.forbes.com/sites/tomdavenport/2020/03/27/return-on-artificial-intelligence-the-challenge-and-the-opportunity/

Forrester research on the challenges and opportunities of operationalizing ML:
"57% of respondents believe silos between data scientists and practitioners inhibit ML deployments."
https://www.capitalone.com/tech/machine-learning/new-forrester-report-on-operationalizing-machine-learning/

On most ML projects never making it into production:
https://towardsdatascience.com/why-production-machine-learning-fails-and-how-to-fix-it-b59616184604

Economics tells the same inauspicious story. The US Census Bureau reported in 2018 that only 2.8% of companies use ML. Meanwhile, a National Bureau of Economic Research paper headed by then-MIT Sloan professor Erik Brynjolfsson identified the "modern productivity paradox." There have been significant advancements in ML technology, but the positive economic results that we would expect to see with such developments are still pending. As the paper puts it, "Systems using AI match or surpass human level performance in more and more domains, leveraging rapid advances in other technologies and driving soaring stock prices. Yet measured productivity growth has declined by half over the past decade."
https://www.nber.org/system/files/working_papers/w24001/w24001.pdf
https://conference.nber.org/conf_papers/f138039.pdf

Ted talk by Erik Brynjolfsson
https://www.ted.com/talks/erik_brynjolfsson_the_key_to_growth_race_with_the_machines/transcript?language=en

Quote from Armin Kakas
https://www.linkedin.com/feed/update/urn:li:activity:6888945163947905024/?commentUrn=urn%3Ali%3Acomment%3A(activity%3A6888945163947905024%2C6892187486450257921)

The top pain points for data teams is "Delivering business impact now through AI."
https://www.expert.ai/resource/harness-the-power-unstructured-data/
https://www.expert.ai/wp-content/uploads/2022/01/harness-the-power-unstructured-data.pdf

"Cultural factors continue to be the greatest obstacle to delivering business value from data investments – The vast majority of data leaders – 79.8% – continue to cite cultural issues — organizational receptivity to change and business transformation, changes to organizational processes, people and skills, organizational alignment, and communications – as the greatest obstacles to realizing business value, reflecting that change is seldom easy, and organizational transformation tends to move slowly. Yet, investment in people skills such as building data literacy, remain low – just 1.6% cited this as their top investment priority. A sign of hope is that 70.9% of CDO/CDAOs and data leaders report that their firms are receptive to change and organizational transformation."

FROM: Data and Analytics Leadership Annual Executive Survey 2023
NewVantage Partners
https://www.newvantage.com/_files/ugd/e5361a_247885043758499ba090f7a5f510cf7c.pdf

Other reasons for failure of data mining
http://www.b-eye-network.com/view/13685

Top 25 Mistakes Corporates Make in their Advanced Analytics Program https://www.datasciencecentral.com/profiles/blogs/top-25-mistakes-corporates-make-in-their-advanced-analytics

Getting started with AI? Start here!
https://hackernoon.com/the-decision-makers-guide-to-starting-ai-72ee0d7044df

12 Predictive Analytics Screw-Ups
https://www.predictiveanalyticsworld.com/machinelearningtimes/12-predictive-analytics-screw-ups/2049/

5 Key Reasons Why Analytics Projects Fail
https://www.elderresearch.com/blog/5-key-reasons-why-analytics-projects-fail/

Getting started with AI? Start here!
https://hackernoon.com/the-decision-makers-guide-to-starting-ai-72ee0d7044df

Choosing the right analytics problem
https://www.elderresearch.com/blog/choosing-the-right-analytics-problem

"Your model is worthless if nobody uses it."
–Mark Tenenholtz, data science entrepreneur and Kaggle Master
https://mobile.twitter.com/marktenenholtz/status/1506964066769637376

CRoss Industry Standard Process for Data Mining (CRISP-DM)
https://en.wikipedia.org/wiki/Cross-industry_standard_process_for_data_mining

For more on CRISP-DM, see the (free PDF) book "9 Laws of Data Mining" by Tom Khabaza, one of its original developers:
https://www.researchgate.net/publication/360159100_9_Laws_of_Data_Mining

The author recommends this 18-minute video covering an ML case study -- stepping end to end through project step -- which complements well the cases in this book. The case is "a prediction tool for electric utilities to be able to predict the level of severity and the locations of power outages in advance of storms impacting their utility territory." (You can view this individual video without registering for the course.)
https://www.coursera.org/learn/managing-machine-learning-projects/lecture/4oWFJ/crisp-dm-case-study

Data and Analytics Leadership Annual Executive Survey 2023
NewVantage Partners
https://www.newvantage.com/_files/ugd/e5361a_247885043758499ba090f7a5f510cf7c.pdf

Why User Education Is Necessary To Avoid AI Failure https://www.forbes.com/sites/naveenjoshi/2022/01/22/why-user-education-is-necessary-to-avoid-ai-failure/

Six Must-Have Mindsets to be a World-Class Data and Analytics Leader
https://iianalytics.com/community/blog/six-must-have-mindsets-to-be-a-world-class-data-and-analytics-leader

"I think there's a misconception that machine learning work is pretty solitary and you can teach yourself to do it or you can do it by yourself on a laptop, but it [takes] a team in order to deploy anything functional that matters... That often gets overlooked because there's a lot of focus on the technology and the right hard skills and the right technical systems. It's really easy to overlook the team dynamics."
- Alyssa Simpson Rochwerger, Blue Shield
Gradient Dissent podcast, May 20, 2021

"There's still an enormous disconnect between what an executive expects to be able to do and... what the machine learning person or the data scientist actually understands as doable."
–James Cham, Partner, Bloomberg Beta
Gradient Dissent, podcast
Episode: James Cham — Investing in the Intersection of Business and Technology
https://www.youtube.com/watch?v=T4LXx8Bs1kY

“We often incorrectly think of deployment of a data science or analytical model as the last stage of the process... Starting with the algorithm first, and only at the end of the project thinking about how to insert it into the business process, is where many deployments fail.”
–Thomas Davenport and Katie Malone, Harvard Data Science Review
Deployment as a Critical Business Data Science Discipline
https://hdsr.mitpress.mit.edu/pub/2fu65ujf/release/3

"One of the most common mistakes [is] failing to plan an implementation at all. You wouldn't think, perhaps, that this is common, but it really is."
–Keith McCormick, machine learning consultant
https://youtu.be/qESQcy_NaN8

Gartner quote:
https://www.gartner.com/en/newsroom/press-releases/2021-03-16-gartner-identifies-top-10-data-and-analytics-technologies-trends-for-2021#:~:text=Most%20analytics%20and%20AI%20projects,of%20analytics%20and%20AI%20assets.

Dean Abbott quote:
https://youtu.be/V7QeNHxuxqY?t=668

"One successful CDO imparted two pieces of advice: 1) Start with a clear connection to business strategy with tangible examples of how data analytics can drive business outcomes

(topline, bottom line, cash, stewardship), and 2) lead with 1-2 forward thinking business partners to demonstrate what is possible. Those partners become the change agents across the organization."
https://hbr.org/2021/08/why-do-chief-data-officers-have-such-short-tenures

Guidance to heed and pitfalls to avoid in the name of achieving model deployment:
https://medium.com/mlwhiz/why-do-machine-learning-projects-fail-9fefb287a66d

On the importance of leadership skills for data science:
https://www.kdnuggets.com/2021/02/fit-lead-data-science.html

Data product management:

Why Your Company Needs Data-Product Managers
https://hbr.org/2022/10/why-your-company-needs-data-product-managers

Why Adopt a Product Orientation for Analytics and AI?
https://iianalytics.com/community/blog/why-adopt-a-product-orientation-for-analytics-and-ai

Machine Learning Operations (MLOps): Overview, Definition, and Architecture
https://arxiv.org/abs/2205.02302

Further reading:

On bridging the divide: "There is a designed-in structural tension between business and data science teams that needs to be recognized and addressed." How? With "high-bandwidth, bidirectional communication channels" between business leaders and data scientists. Assign an "innovation marshal". There's a "full-frontal attack and disruption signaled by" data science.
https://sloanreview.mit.edu/article/to-succeed-with-data-science-first-build-the-bridge/

Nice article on the importance of leadership per se, relative to "hands-on":
https://www.kdnuggets.com/2021/02/fit-lead-data-science.html

On academic programs in analytics not keeping up with leadership skills:
https://pubsonline.informs.org/doi/10.1287/inte.2018.0955

Overview on the executive leadership requirements for data science:
https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/ten-red-flags-signaling-your-analytics-program-will-fail

"At the end of the day, data scientists are change agents. If nothing changes as a result of what your model does, then you didn't really do your job."
–Jodi Blomberg Senior, Director, Data Science and Machine Learning, Waste Management, Inc.

Words of wisdom on managing ML projects:

"Getting started with AI? Start here!"
https://hackernoon.com/the-decision-makers-guide-to-starting-ai-72ee0d7044df

"Choosing the right analytics problem"
https://www.elderresearch.com/blog/choosing-the-right-analytics-problem/

ML newcomer pitfalls:
https://www.reddit.com/r/datascience/comments/vlpi4u/what_are_the_most_common_mistakes_you_see_junior/

Anaconda's 2021 "State of Data Science" survey
https://know.anaconda.com/rs/387-XNW-688/images/Anaconda-2021-SODS-Report-Final.pdf