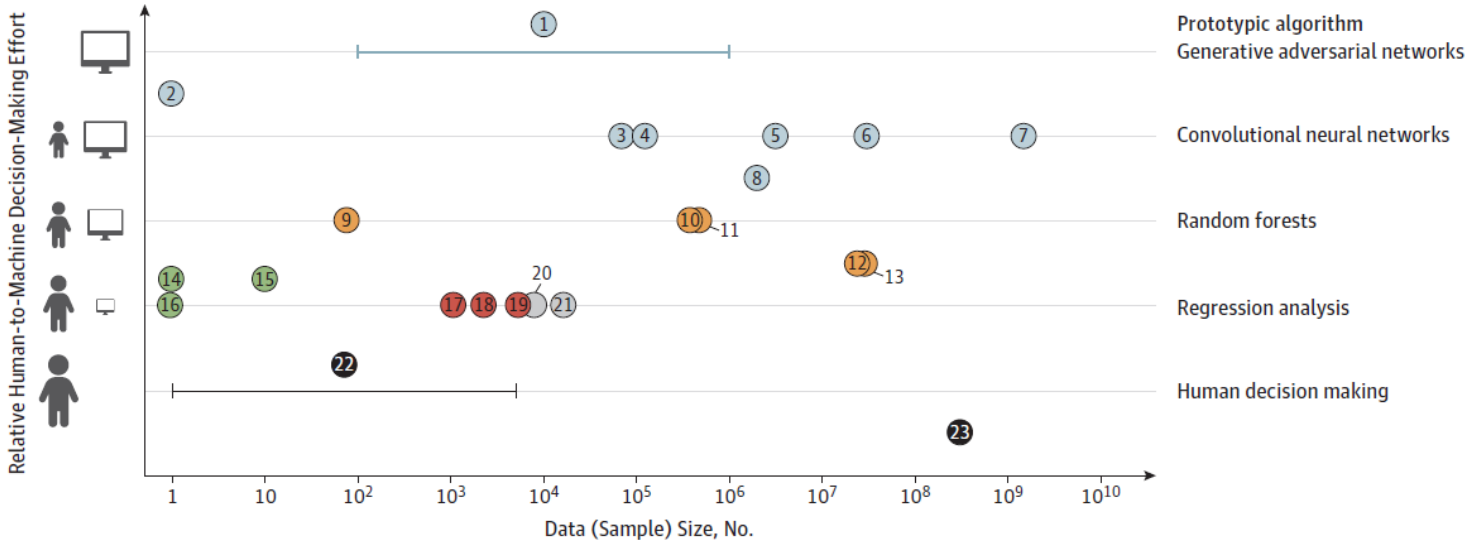
Supplementary Online Content

Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA*. doi:10.1001/jama.2017.18391

This supplementary material has been provided by the authors to give readers additional information about their work.

eFigure. Illustrative Examples of Machine Learning



Deep learning

1. Generative adversarial networks (2014)
2. Google AlphaGo Zero (2017)
3. ATM check readers (1998)
4. Google diabetic retinopathy (2016)
5. ImageNet computer vision models (2012-2017)
6. Google AlphaGo (2015)
7. Facebook Photo Tagger (2015)
8. Prediction of 1-y all-cause mortality (2017)

Classic machine learning

9. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling (2002)
10. EHR-based CV risk prediction (2017)
11. Netflix Prize winner (2006)
12. Google Search (1998)
13. Amazon product recommendation (2003)

Expert AI systems

14. MYCIN (1975)
15. CASNET (1982)
16. DXplain (1986)

Risk calculators

17. CHA_2DS_2-VASc Score for atrial fibrillation stroke risk (2017)
18. MELD end-stage liver disease risk score (2001)
19. Framingham CV risk score (1998)

Randomized Clinical Trials

20. Celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis (2002)
21. Use of estrogen plus progestin in healthy postmenopausal women (2002)

Other

22. Clinical wisdom
23. Mortality rate estimates from US Census (2010)

Abbreviations: ATM, automated teller machine; CV, cardiovascular; EHR, electronic health record; MELD, model for end-stage liver disease.

Prototypical algorithms

Generative adversarial networks: Neural networks that learn everything from data with no human input. These networks learn complex statistical patterns in the data that can be used to generate new synthetic data points that look like real data.

Convolutional neural networks: Neural networks that learn directly from raw images, which allows for discovery of complex and nuanced relationships not known beforehand.

Random forests: A collection of "if-then-else" statements learned from data. Allows for nonlinear variable interactions to be automatically discovered.

Regression analysis: A statistical model that assumes a specified relationship between the dependent and independent variables. Popular choices include linear and logistic regression models.

Human decision making: Decisions that are made without formal analyses of data, such as wisdom, experience, intuition, and heuristics.

eReferences.

		Year of Creation	Reference
	Deep Learning		
①	Generative adversarial networks	2014	Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. *Adv Neural Information Processing Syst*. 2014:2672-2680.
②	Google AlphaGo Zero	2017	Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of Go without human knowledge. *Nature*. 2017;550(7676):354-359.
③	ATM check readers	1998	LeCun Y, et al. Handwritten digit recognition with a back-propagation network. *Adv Neural Inf Process Syst*. 1990.
④	Google diabetic retinopathy	2016	Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402-2410.
⑤	ImageNet computer vision models	2012-2017	Russakovsky O, et al. Imagenet large scale visual recognition challenge. *Int J Comput Vis*. 2015;115(3):211-252.
⑥	Google AlphaGo	2015	Silver D, Huang A, Maddison CJ, et al. Mastering the game of Go with deep neural networks and tree search. *Nature*. 2016;529(7587):484-489.
⑦	Facebook Photo Tagger	2015	Inside Facebook's biggest artificial intelligence project ever. http://fortune.com/facebook-machine-learning/.
⑧	Prediction of 1-year all-cause mortality	2017	Avati A, et al. Improving palliative care with deep learning. arXiv preprint. 2017;1711.06402.
	Classic Machine Learning		
⑨	Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling	2002	Shipp MA, Ross KN, Tamayo P, et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med*. 2002;8(1):68-74.
⑩	EHR-based CV risk prediction	2017	Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PloS One*. 2017;12(4):e0174944.
⑪	Netflix Prize winner	2006	Bell RM, Koren Y. Lessons from the Netflix prize challenge. *Acm Sigkdd Explorations Newsletter*. 2007;9(2):75-79.
⑫	Google Search	1998	Page L, et al. The PageRank citation ranking: bringing order to the web. Stanford InfoLab; 1999.
⑬	Amazon product recomendation	2003	Linden G, Smith B, York J. Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing*. 2003;7(1):76-80.
	Expert AI Systems		
⑭	MYCIN	1975	Shortliffe E, ed. *Computer-Based Medical Consultations: MYCIN*. Vol 2.

			Elsevier; 2012.
⑮	CASNET	1982	Kulikowski CA, Weiss SM. Representation of expert knowledge for consultation: the CASNET and EXPERT projects. *Artif Intell Med*. 1982;51.
⑯	DXplain	1986	Barnett GO, Cimino JJ, Hupp JA, Hoffer EP. DXplain: an evolving diagnostic decision-support system. *JAMA*. 1987;258(1):67-74.
	Risk Score		
⑰	CHA_2DS_2-VASc score for atrial fibrillation stroke risk	2017	Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the Euro heart survey on atrial fibrillation. *Chest*. 2010;137(2):263-272.
⑱	MELD end-stage liver disease risk score	2001	Kamath PS, Wiesner RH, Malinchoc M, et al. A model to predict survival in patients with end-stage liver disease. *Hepatology*. 2001;33(2):464-470.
⑲	Framingham CV risk score	1998	Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97(18):1837-1847.
	Randomized Clinical Trial		
⑳	Celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis	2002	Silverstein FE, Faich G, Goldstein JL, et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomized controlled trial. *JAMA*. 2000;284(10):1247-1255.
㉑	Use of estrogen plus progestin in healthy postmenopausal women	2002	Rossouw JE, Anderson GL, Prentice RL, et al; Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288(3):321-333.
	Other		
㉓	Mortality rate estimate from US Census	2010	US Census