

Name	Year	Source for year	Don't include?	
Matrices	200	The Nine Chapters on the Mathematical Art		
Gaussian elimination	200	"The Nine Chapters on the Mathematical Art"		
Constraint satisfaction/solvers	300	Diophantus, Arithmetica		
Median	1250	https://en.wikipedia.org/wiki/Median#History		
Nearest neighbor	1644	http://mathworld.wolfram.com/VoronoiDiagram.html		
Decision theory	1654	Pascal		
Discrete probability	1657	https://en.wikipedia.org/wiki/History_of_probability		
Newton's method	1669	Miracle Years		
Differential calculus	1669	Newton		
Integral calculus	1669	Newton		
Calculus	1671	https://en.wikipedia.org/wiki/History_of_calculus		
system of ODEs	1671	Newton, Isaac. (c.1671). Methodus Fluxionum et Serierum Infinitarum (The Method of Fluxions and Infinite Series), published in 1736 [Opuscula, 1744, Vol. I. p. 66].		
Convolution	1754	D'Alembert		
Calculus of Variations	1756	https://en.wikipedia.org/wiki/Calculus_of_variations#History		
Bayes	1763	https://en.wikipedia.org/wiki/Bayes%27_theorem		
Euler's method	1768	Institutionum calculi integralis		
Monte Carlo	1777	https://en.wikipedia.org/wiki/Buffon%27s_needle		
linear regression	1805	https://en.wikipedia.org/wiki/Least_squares		
QR decomposition	1820	Laplace		
Standard deviation	1821	http://jeff560.tripod.com/m.html (mean error)		
Combinatorial optimization	1825	Hamilton and Kirkman		
Compression	1838	Morse code		
Pure strategy Nash equilibrium	1838	https://en.wikipedia.org/wiki/Cournot_competition		
Parallel programming	1842	Menabrea, L. F. (1842). Sketch of the Analytic Engine Invented by Charles Babbage.		
Graph Laplacian	1847	https://arxiv.org/pdf/1111.2897.pdf		

Name	Year	Source for year	Don't include?
Differential geometry	1857	Riemann	
stochastic process	1859	Truesdell, C. (1975). "Early kinetic theories of gases". Archive for History of Exact Sciences. 15 (1): 31–32. doi:10.1007/BF00327232. ISSN 0003-9519.	
Natural selection	1859	Darwin	
Poisson models	1860	https://en.wikipedia.org/wiki/Poisson_point_process#History	
Boolean algebra	1864	https://books.google.com/books?id=-AokWWhbLUIC&pg=PA2#v=onepage&q&f=false	
DFS	1876	https://en.wikipedia.org/wiki/Depth-first_search#cite_note-2	
correlation	1885	Rodgers, J. L.; Nicewander, W. A. (1988). "Thirteen ways to look at the correlation coefficient". The American Statistician. 42 (1): 59–66. doi:10.1080/00031305.1988.10475524. JSTOR 2685263.	
dynamical systems	1892	"New Methods of Celestial Mechanics", Poincare	
Runge Kutta	1895	Runge, Carl David Tolmé (1895), "Über die numerische Auflösung von Differentialgleichungen", Mathematische Annalen, Springer, 46 (2): 167–178, doi:10.1007/BF01446807.	
chaos theory	1898	Hadamard, Jacques (1898). "Les surfaces à courbures opposées et leurs lignes géodesiques". Journal de Mathématiques Pures et Appliquées. 4: 27–73.	
PCA	1901	https://en.wikipedia.org/wiki/Principal_component_analysis	
Representation learning	1901	https://www.sciencedirect.com/science/article/pii/S2405918816300459	
Central limit theorem	1901	https://en.wikipedia.org/wiki/Central_limit_theorem#History	
Markov Chains	1906	https://en.wikipedia.org/wiki/Markov_chain#History	
Spiking neural nets	1907	Abbott, L.F. (1999). "Lapique's introduction of the integrate-and-fire model neuron (1907)" (PDF). Brain Research Bulletin. 50 (5/6): 303–304. doi:10.1016/S0361-9230(99)00161-6. PMID 10643408. Archived from the original (PDF) on 2007-06-13. Retrieved 2007-11-24.	
Operant Conditioning	1911	Thorndike, E. L. (1911) Animal intelligence: Experimental studies. Macmillan	
ZF set theory	1917	https://en.wikipedia.org/wiki/Zermelo%E2%80%93Fraenkel_set_theory#History	
First-order logic	1917	https://philosophy.stackexchange.com/questions/2617/how-did-first-order-logic-come-to-be-the-dominant-formal-logic	

Name	Year	Source for year	Don't include?	
Signal/noise distinction	1918	Schottky, W. (1918). "Über spontane Stromschwankungen in verschiedenen Elektrizitätsleitern". Annalen der Physik (in German). 57: 541–567. Bibcode: 1918AnP...362..541S. doi:10.1002/andp.19183622304.		
Causal graphs	1921	https://en.wikipedia.org/wiki/Causal_graph		
Rate coding	1926	Adrian ED, Zotterman Y (1926). "The impulses produced by sensory nerve endings: Part II: The response of a single end organ". J Physiol. 61: 151–171. doi:10.1113/jphysiol.1926.sp002281. PMC 1514782 Freely accessible.[1]		
Classical Conditioning	1927	Pavlov, I. (1927). Conditioned reflexes. London: Oxford University Press		
Power methods	1927	von Mises		
Theorem prover	1928	https://en.wikipedia.org/wiki/Entscheidungsproblem		
Godel's Incompleteness	1931			
Thompson sampling	1933	https://en.wikipedia.org/wiki/Thompson_sampling		
Measure theoretic probability	1933	Grundbegriffe der Wahrscheinlichkeitsrechnung		
Sparse matrices in ML	1934	"Latent class analysis as an association model for information retrieval"		
Exponential family	1935	https://en.wikipedia.org/wiki/Exponential_family		
Turing machine/universality	1936	https://en.wikipedia.org/wiki/Turing_machine		
Lambda calculus	1936	Church Paper		
Electronic computers	1937	Atanasoff–Berry computer		
linear programming	1939	https://en.wikipedia.org/wiki/Linear_programming		
Finite state automata	1943	"A Logical Calculus Immanent in Nervous Activity"		
Mixed strategy Nash equilibrium	1944	https://en.wikipedia.org/wiki/Nash_equilibrium#History		
BFS	1945	https://en.wikipedia.org/wiki/Breadth-first_search		
Fast sorting algorithms (n log n)	1945	https://en.wikipedia.org/wiki/Merge_sort		
Von Neumann architecture	1945	First Draft of a Report on the EDVAC.		
MCMC	1946	https://en.wikipedia.org/wiki/Monte_Carlo_method#History		
Cox's theorem	1946	https://aapt.scitation.org/doi/10.1119/1.1990764		
simplex algorithm	1947	https://en.wikipedia.org/wiki/Simplex_algorithm		
VNM Utility	1947	Neumann, John von and Morgenstern, Oskar, Theory of Games and Economic Behavior		
Bits (to measure information)	1948	https://en.wikipedia.org/wiki/Bit		

Name	Year	Source for year	Don't include?
Entropy	1948	https://en.wikipedia.org/wiki/Entropy_(information_theory)	
Cybernetics	1948	Wiener, Cybernetics, or Control and Communication in the Animal and the Machine	
Hebbian learning	1949	https://en.wikipedia.org/wiki/Hebbian_theory	
Machine translation	1949	Machine translation	
MiniMax	1950	http://stanford.edu/~cpiech/cs221/apps/deepBlue.html	
Backtracking search	1950	https://books.google.com/books?id=Kjap9ZWcKOoC&pg=PA14	
Center surround	1950	Neurophysiology	
Gabor filter	1950		
KL Divergence / cross entropy	1951	https://en.wikipedia.org/wiki/Kullback%E2%80%93Leibler_divergence#Etymology	
Importance sampling	1951	https://stats.stackexchange.com/questions/116601/when-was-importance-sampling-first-stated	
Compilers	1951	Corrado Böhm's PhD thesis	
LISP	1952	https://en.wikipedia.org/wiki/Lisp_(programming_language)	
K-armed bandit	1952	https://en.wikipedia.org/wiki/Multi-armed_bandit	
Speech recognition	1952	https://web.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/354_LALI-ASRHistory-final-10-8.pdf	
Metropolis hastings (MCMC)	1953	https://arxiv.org/pdf/0808.2902.pdf	
Genetic algorithms	1954	https://en.wikipedia.org/wiki/Genetic_algorithm#History	
K-Means	1955	https://pdfs.semanticscholar.org/1043/3d2bdf7fd0d6fb9135cc8aa111b4eb6e7d91.pdf	
quadratic programming	1956	https://link.springer.com/chapter/10.1007/978-0-387-77439-8_8	
Stochastic gradient descent	1957	https://en.wikipedia.org/wiki/Perceptron	
Peceptron	1957	https://en.wikipedia.org/wiki/Perceptron	
MDP	1957	Bellman, R. (1957). "A Markovian Decision Process". <i>Journal of Mathematics and Mechanics</i> . 6.	
Bounded rationality	1957	<i>Models of Man</i>	
Dynamic Programming	1957	Bellman, R. E. (1957). <i>Dynamic Programming</i> . Princeton University Press, Princeton, NJ.	

Name	Year	Source for year	Don't include?
Curse of dimensionality	1957	https://en.wikipedia.org/wiki/Curse_of_dimensionality#cite_note-1	
logistic regression	1958	https://en.wikipedia.org/wiki/Logistic_regression	
Symbolic AI	1959	https://en.wikipedia.org/wiki/General_Problem_Solver	
Hidden Markov Model	1960	https://en.wikipedia.org/wiki/Hidden_Markov_model#History	
Exploratory factor analysis	1960	https://en.wikipedia.org/wiki/Exploratory_factor_analysis	
Functionalism (theory of mind)	1960	https://plato.stanford.edu/entries/functionalist/#AntFun (Putnam)	
RL	1960	Howard, Dynamic Programming and Markov Processes	
Multilayer perceptron	1960		
Conjugate prior	1961	Howard Raiffa and Robert Schlaifer. Applied Statistical Decision Theory	
L2 Norm regularization	1962	On the stability of inverse problems, Tikhonov	
K-Complexity	1963	https://en.wikipedia.org/wiki/Kolmogorov_complexity	
Alpha beta pruning	1963	https://en.wikipedia.org/wiki/Alpha%E2%80%93beta_pruning	
SVM	1963	https://en.wikipedia.org/wiki/Support_vector_machine	
Simulated annealing for optimization	1963	CD Gelatt, and MP Vecchi. Optimisation by simulated annealing	
Solomonoff Induction	1964	https://en.wikipedia.org/wiki/Solomonoff%27s_theory_of_inductive_inference	
Kernel methods	1964	https://en.wikipedia.org/wiki/Kernel_perceptron	
FFT	1965	https://en.wikipedia.org/wiki/Fast_Fourier_transform	
Knowledge based systems	1965	http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.34.9207&rep=rep1&type=pdf	
POMDP	1965	Åström, K.J. (1965). "Optimal control of Markov processes with incomplete state information". Journal of Mathematical Analysis and Applications. 10: 174–205. doi:10.1016/0022-247X(65)90154-X.	
particle filter	1965	https://en.wikipedia.org/wiki/Particle_filter#cite_note-dm962-1	
Image recognition	1966	Papert, Seymour (1966-07-01). "The Summer Vision Project". MIT AI Memos (1959 - 2004). Retrieved 2 August 2016.	
Cross validation	1968	https://stats.stackexchange.com/questions/26696/who-invented-k-fold-cross-validation	

Name	Year	Source for year	Don't include?
Minimum message length	1968	"An information measure for classification" Wallace & Boulton	
Bidirectional search	1969	Pohl, Ira. Bidirectional and heuristic search in path problems. No. SLAC-R-104. 1969.	
Frame problem	1969	https://en.wikipedia.org/wiki/Frame_problem	
conceptual dependency model	1969	Roger Schank, 1969, A conceptual dependency parser for natural language Proceedings of the 1969 conference on Computational linguistics, Sång-Säby, Sweden pages 1-3	
Newcomb's Problem	1969	https://en.wikipedia.org/wiki/Newcomb%27s_paradox	
Test and validation set	1970		
Automatic differentiation	1970	https://en.wikipedia.org/wiki/Backpropagation#History	
P vs NP	1971	https://en.wikipedia.org/wiki/P_versus_NP_problem#History	
n-gram models	1971	https://books.google.com/books/about/Markov_models_and_linguistic_theory.html?id=qxBZAAAAMAAJ	
VC dimension	1971	https://en.wikipedia.org/wiki/VC_dimension	
Place cells	1971	O'Keefe and Dostrovsky	
Prolog	1972	https://en.wikipedia.org/wiki/Prolog#History	
Dataset augmentation	1974	Allen, David M. "The relationship between variable selection and data agumentation and a method for prediction." Technometrics 16, no. 1 (1974): 125-127.	
correlated equilibrium	1974	https://en.wikipedia.org/wiki/Correlated_equilibrium	
BPP	1977	https://www.sciencedirect.com/science/article/pii/S0019995886800444	
Hyperparameters	1979	First instance I could find on Google Scholar, Akaike, Hirotugu. "Ignorance prior distribution of a hyperparameter and Stein's estimator." Annals of the Institute of Statistical Mathematics 32.1 (1980): 171-178.	
CDT	1981	https://en.wikipedia.org/wiki/Causal_decision_theory#cite_note-13	
RNN / Hopfield network	1982	https://en.wikipedia.org/wiki/Recurrent_neural_network#History	
Back propagation	1982	https://en.wikipedia.org/wiki/Backpropagation	
parasitic attractor	1982	Hopfield, J. J. 1982. Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. U.S.A. 79, 2554-2558.	

Name	Year	Source for year	Don't include?
Actor-Critic	1983	Barto, Sutton, and Anderson, 1983	
Anthropics	1983	The Doomsday Argument, The anthropic principle and its implications for biological evolution	
program equilibrium	1984	https://intelligence.org/files/ProgramEquilibrium.pdf	
ICA	1984	https://en.wikipedia.org/wiki/Independent_component_analysis#History_and_background	
System 1 / system 2	1984	Evans, Jonathan (1984). "Heuristic and analytic processes in reasoning". British Journal of Psychology. 75: 451–468. doi:10.1111/j.2044-8295.1984.tb01915.x.	
Temporal difference learning	1984	http://www.scholarpedia.org/article/Temporal_difference_learning#History	
Bayesian network	1985	https://en.wikipedia.org/wiki/Bayesian_network#History	
Sparse coding	1985	General neuro and ML literature	
Boltzmann Machine	1985	Hinton	
Momentum	1986	https://en.wikipedia.org/wiki/Stochastic_gradient_descent#Momentum	
Restricted Boltzmann machine	1986	https://en.wikipedia.org/wiki/Restricted_Boltzmann_machine	
Lojban	1987	https://en.wikipedia.org/wiki/Lojban	
MCTS	1987	https://en.wikipedia.org/wiki/Monte_Carlo_tree_search#History	
Autoencoders	1987	https://stats.stackexchange.com/questions/238381/what-is-the-origin-of-the-autoencoder-neural-networks	
Variational Bayes	1987	https://arxiv.org/pdf/1601.00670.pdf	
Policy gradients	1987	https://en.wikipedia.org/wiki/Reinforcement_learning#Direct_policy_search	
Catastrophic interference	1989	McCloskey, M. & Cohen, N. (1989) Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (ed.) The Psychology of Learning and Motivation, 24, 109-164	
Q learning	1989	https://en.wikipedia.org/wiki/Q-learning	
Universal function approximator (proof for neu	1989	https://en.wikipedia.org/wiki/Universal_approximation_theorem	
Spectral clustering	1990	https://arxiv.org/pdf/0711.0189.pdf	
Minibatch	1992	Moller, Martin. "Supervised learning on large redundant training sets." In Neural Networks for Signal Processing [1992] II., Proceedings of the 1992 IEEE-SP Workshop, pp. 79-89. IEEE, 1992.	
Self-indication assumption	1992	https://en.wikipedia.org/wiki/Self-indication_assumption	

Name	Year	Source for year	Don't include?
SARSA	1994	http://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf	
Interior point method applied to convex optimiz	1994	https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf	
Random forests	1995	https://en.wikipedia.org/wiki/Random_forest	
Soft margin SVM	1995	https://en.wikipedia.org/wiki/Support_vector_machine#History	
L1 Norm regularization	1996	https://en.wikipedia.org/wiki/Lasso_(statistics)	
Gated memory (LSTM, GRU)	1997	https://en.wikipedia.org/wiki/Long_short-term_memory	
No Free Lunch Theorem	1997	https://en.wikipedia.org/wiki/No_free_lunch_theorem	
ConvNets	1998	https://en.wikipedia.org/wiki/Convolutional_neural_network#History	
Hierarchical RL	1998	https://people.cs.umass.edu/~mahadeva/papers/hrl.pdf	
Locality sensitive hashing	1998	Indyk, Piotr, and Rajeev Motwani. "Approximate nearest neighbors: towards removing the curse of dimensionality." In Proceedings of the thirtieth annual ACM symposium on Theory of computing, pp. 604–613. ACM, 1998.	
Information bottleneck	1999	Tishby, Naftali; Pereira, Fernando C.; Bialek, William (September 1999). The Information Bottleneck Method (PDF). The 37th annual Allerton Conference on Communication, Control, and Computing. pp. 368–377.	
AIXI	2000	https://en.wikipedia.org/wiki/AIXI	
AIXI-TL	2000	https://arxiv.org/pdf/cs/0004001.pdf	
LDA	2000	https://en.wikipedia.org/wiki/Latent_Dirichlet_allocation	
SMT solvers	2001	https://escape.cis.upenn.edu/documents/ClarkBarrettSlides.pdf	
Rademacher and Gaussian complexity	2002	http://www.jmlr.org/papers/volume3/bartlett02a/bartlett02a.pdf	
Pooling (deep learning)	2002	Walther, Dirk, et al. "Attentional selection for object recognition—a gentle way." International Workshop on Biologically Motivated Computer Vision. Springer, Berlin, Heidelberg, 2002.	
Neural word embeddings	2003	http://www.jmlr.org/papers/volume3/bengio03a/bengio03a.pdf	
Variational message passing	2005	http://www.jmlr.org/papers/volume6/winn05a/winn05a.pdf	
Q learning + ML	2005	http://ml.informatik.uni-freiburg.de/former/_media/publications/riecml05.pdf	
Church probabilistic programming language	2008	https://web.stanford.edu/~ngoodman/papers/churchUAI08_rev2.pdf	

Name	Year	Source for year	Don't include?
Group theoretic machine learning	2008	Risi Kondor's paper	
Convergent instrumental goals	2008	Omohundro, Stephen M. "The basic AI drives." AGI. Vol. 171. 2008.	
High dimensional computing	2009	Kanerva, Pentti (2009) Hyperdimensional Computing: An Introduction to Computing in Distributed Representation with High-Dimensional Random Vectors, Cognitive Computation, Volume 1, Issue 2, pp. 139–159.	
UDT	2009	https://www.lesswrong.com/posts/de3xjFaACCAk6imzv/towards-a-new-decision-theory	
Attention models	2010	http://www.wildml.com/2016/01/attention-and-memory-in-deep-learning-and-nlp/	
Xavier initialization	2010	http://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf	
online convex optimization	2011	https://www.cs.huji.ac.il/~shais/papers/OLsurvey.pdf	
Variational autoencoders	2013	https://arxiv.org/abs/1312.6114	
Deep Q learning	2013	https://arxiv.org/abs/1312.5602	
Drop out	2014	http://jmlr.org/papers/v15/srivastava14a.html	
GAN	2014	https://en.wikipedia.org/wiki/Generative_adversarial_network	
Reflective oracles	2015	https://arxiv.org/abs/1508.04145	
Adversarial training	2015	https://arxiv.org/pdf/1412.6572.pdf	
Batch normalization	2015	Ioffe, S. and Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167.	
Logical Induction	2016	https://intelligence.org/2016/09/12/new-paper-logical-induction/	
Asynchronous actor-critic	2016	http://proceedings.mlr.press/v48/mniha16.pdf	
Binarized neural nets	2016	https://arxiv.org/abs/1602.02830	
Capability amplification	2016	https://ai-alignment.com/policy-amplification-6a70cbee4f34	
Cycle GAN	2017	https://arxiv.org/abs/1703.10593	
OpenAI evolution thing	2017	https://blog.openai.com/evolution-strategies/	
topological sort	?		
Supervised vs unsupervised learning	?		
Adversarial examples	?		
Propositional logic	250 BC	https://en.wikipedia.org/wiki/Propositional_calculus	

Name	Year	Source for year	Don't include?	
Robotics	300 BC	<i>Liezi</i> automata		
Aristotelian logic	350 BC	https://en.wikipedia.org/wiki/Prior_Analytics		
Occam's razor	350 BC	https://en.wikipedia.org/wiki/Occam%27s_razor#History		
Orthogonality thesis	425 BC	Herodotus, "The Histories"		
Materialism	600 BC	https://en.wikipedia.org/wiki/Materialism		
Skip layer	n/a		Yes (trivial from NNs)	